A Modelling Approach For Evaluating The Ranking Capability of Situational Awareness System In Real Time Operation

Modelling, evaluating and quantifying different situational assessment in real time operation, using an analytical approach for measuring the ranking capability of SWA system

Orabi Mahmoud Fahmi SHURRAB

Faculty of Engineering and Informatics

University of Bradford

Submitted for the Degree of

Doctor of Philosophy

2016

Abstract

Orabi Mahmoud Fahmi Shurrab

A Modelling Approach for Evaluating the Ranking Capability of Situational Awareness System in Real Time Operation

Modelling, Evaluating and Quantifying Different Situational Assessment in Real Time Operation, Using an Analytical Approach for Measuring the Ranking Capability of SWA System

Keywords: Situational Awareness, Performance Metrics, Data Fusion, Operational Research, Process Refinement

In a dynamically monitored environment the analyst team need timely and accurate information to conduct proactive action over complex situations. Typically, there are thousands of reported activities in a real time operation, therefore steps are taken to direct the analyst's attention to the most important activity. The data fusion community have introduced the information fusion model, with multiple situational assessments. Each process lends itself to ranking the most important activities into a predetermined order.

Unfortunately, the capability of a real time system can be hindered by the knowledge limitation problem, particularly when the underlying system is processing multiple sensor information. Consequently, the situational awareness domains may not rank the identified situation as perfect, as desired by the decision-making resources. This

thesis presents advanced research carried out to evaluate the ranking capability of information from the situational awareness domains: perception, comprehension and projection. The Ranking Capability Score (RCS) has been designed for evaluating the prioritisation process. The enhanced (RCS) has been designed for addressing the knowledge representation problem in the user system relation under a situational assessment where the proposed number of tracking activities are dynamically shifted. Finally, the Scheduling Capability Score was designed for evaluating the scheduling capability of the situational awareness system.

The proposed performance metrics have been successful in fulfilling their objectives. Furthermore, they have been validated and evaluated using an analytical approach, through conducting a rigorous analysis of the prioritisation and scheduling processes, despite any constraints related to a domain-specific configuration.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text. Part of this research has been peer reviewed and published in the following work:

Publication

1. [79] Shurrab, O. and Awan, I. (2015). Measuring the ranking capability of SWA system. In Future Internet of Things and Cloud (FiCloud), 2015 3rd International Conference on, pages 165–172.

2. [80] Shurrab, O. and Awan, I. (2015). Performance evaluation for process refinement stage of SWA system. In Future Internet of Things and Cloud (FiCloud), 2015 3rd International Conference on, pages 240–247. IEEE.

3. [78] Shurrab, O. (2016) Responding to emerging situation by developing the ranking capability score (rcs). In Multisensor Fusion and Integration for Intelligent Systems (MFI), the 16th international Conference pages 582-587.

4. [81] Shurrab, O. (2016) Toward an optimisation technique for dynamically monitored environment. SPIE Remote Sensing, International Society for Optics and Photonics. Pages 100070G-100070G.

Orabi Mahmoud Fahmi SHURRAB

2016

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Overviews

There are five different domains where the analyst's team must keep up to date with a dynamically monitored environment. These domains are classified as sea, land, space, air and cyberspaces. To meet the needs of these domains, the data fusion community have introduced the information fusion reference model[84], [10],[26], [27] [11],[14],[7], [8],[12], [13], [33],[34] for multi-disciplinary areas, which describes the theoretical concept of a real-time system with multiple levels of situational assessment. Each level performs a contextual task during a real time operation. In return, the proposed output of these simultaneous processes are either prioritised in a list of events (namely the tracking activity), or represented by visual means, supporting timely responses during a real time operation.

The first three levels of assessment are regarded as the core processes of a situational awareness (SWA) system. The first level of assessment (perception stage) is an automated process based on a predefined configuration. The purpose of this level is to assess and organise multiple sensor information into high abstract views for any emerging situation. Simultaneously, the next two levels of assessment, level 2

(comprehension stage) and level 3 (projection stage), are also automated processes which rank identified tracking activities, based either on their current state or anticipated developments from the perceived situations. This is achieved through consulting different pieces of information concerning the dynamically monitored environment.

Unfortunately, the capability of a SWA system can be hindered by the knowledge limitation problem, specifically when the underlying system is processing multiple sensor information during a real time operation. Consequently the system may not rank the identified list of tracking activities as perfect, as desired by the decision-making resources. With this in mind, researchers have defined two further advanced levels for the real time system; the first level is the process refinement (level 4) for evaluating the performance of the real time system and the second is the user refinement (level 5) for addressing knowledge representation issues, concerning the user system relation.

This section has discussed the theoretical concept of a real time system and the multiple levels of situational assessments. The next section discusses the motivation of this thesis.

1.2 Motivations

The process refinement(level 4) of the Joint Directors of Laboratories (JDL) is a meta-process used to assess and improve the data fusion task for supporting decision-making resources during a real time operation. During the assessment stage, the underlying process is expected to assess the performance of a real time system. Furthermore, the verification techniques contain two forms of evaluation; the first method is the qualitative approach and the second method is quantitative.

Hence, during the qualitative stage, researchers [19][20] [50] [7] [11][7] [77] have developed a number of methods to investigate the capability of a real time system. However, it requires predefined knowledge in order to serve only a domain specific

configuration. On the other hand, during the quantitative assessment, [15] [4] [90][76] [77], the evaluation method is capable of assessing different domains with minimum or no predefined knowledge about that specific domain; such a method can serve wider views in comparison to the qualitative evaluation method. Therefore, this thesis intends to develop advanced quantitative methods for evaluating the performance of a SWA system.

Furthermore, Salerno[74] has proposed a scoring scheme to evaluate the capability of SWA systems, in terms of ranking important events into a contextual order. Tadda [89] explained that the activities of interest (AoI) are far from the top of the lists, thus analysts take longer in assessing such situations. The scoring techniques provide an indication of how close the analysts are to the most important activity. Unfortunately, the AOI scores are limited to contextual situational assessment. Blasch [15] advocated for an extension to have wider views for the AoI scores.

According to our best knowledge, less attention has been given to the performance evaluation with regards to the ranking capability of a real time system. Originally, existing performance metrics [15], [37] [76],[37], [74], [77] [89] [90] had not been designed for measuring the ranking capability of the SWA system. Specifically, corner cases of different situational assessment needs and configurations have not been considered. This thesis presents advanced research work implemented to evaluate the ranking capability of an SWA system for a number of different scenarios.

1.3 Aims and Objectives

The aim of this project is to propose a quantitative assessment method for evaluating the ranking capability of a real time system to serve a wide number of domains as well as different situational assessment needs and configurations.

The aim can be achieved through the following objectives:

1-Deriving a mechanisms for evaluating the prioritisation and scheduling processes of a real time system to quantify the information perception, comprehension and projection processes. The proposed performance metric intends to evaluate the capability of a real time system using quantitative assessment to serve a wider number of domains during the real time operation.

2-Implementing a quality based evaluation techniques for verifying the proposed performance metrics against its intended purpose. The underpinning evaluation will encompass three phases. The first phase will use an analytical approach to compute the number of ranking instances for any given scenarios concerning the prioritisation process or scheduling process of a real time system. The second phase will use Matlab to simulate all the ranking instances computed during the analytical stage. The third phase intends to examine the potential of the proposed scoring scheme in terms of providing a unique score for all possible ranking instances proposed by the simulation phase.

3-Developing a reliability based evaluation for verifying the proposed performance metric against the decision-making perception. This is to address the knowledge representation problem concerning system-user perceptions when the real time system is experiencing ranking capability issues in a dynamically monitored environment.

4-Defining two modelling schemes for representing information perception, comprehension and projection in the form of listing tracking activities. The first modelling scheme will be designed to evaluate the prioritisation process of a real time system and the second modelling scheme will be designed to evaluate the scheduling process of a multi level situation assessment.

5-The proposed performance metrics will be designed and evaluated using an analytical approach. Such methods will allow the evaluation process to conduct a

rigorous analysis of the prioritisation and scheduling processes, despite any constraints related to a domain-specific configuration.

6-Driving a mechanism to analyse the computational complexity issues for two different operations involved in evaluating the prioritisation and scheduling processes for a real time system. The first operation will occur during the assessment stage where the underlying performance metric is required to compute only essential values for assessing the capability of a real time system. The second operation will occur when the potential performance metric is required to compute more values to assess the optimisation technique concerning the ranking capability of the SWA system.

1.4 Thesis Organisation

In Chapter 2, the thesis discusses the theoretical concept of a real time system by reviewing a relevant reference model which has been designed to support a real time operation. The next section will discuss literature reviews and highlight four different research problems concerned with the capability of the SWA system. The first research problem concerns knowledge based issues, the second concerns ranking capability issues, the third concerns the process refinement stage for assessing the performance of real time systems and the fourth concerns knowledge representation issues regarding user perception. Furthermore, as this thesis is focused on evaluating the ranking capability of real time systems, the last section will further discuss the theoretical concept of ranking, prioritisation and scheduling processes from the information fusion perspective. This is to explain how each process is likely to occur during multi level situational assessments.

Chapter 3 presents two contributions. The first develops a modelling scheme for representing the outputs of a SWA system in the form of a list of prioritised events. The second contribution introduces the "Ranking Capability Score" (RCS) as well as

a guidance case study for evaluating the ranking capability of a SWA system. This will, primarily, deal with the prioritisation process of a real time system, under a contextual scenario where the SWA system has identified only a number of tracking activities regarded as important, but each with a different degree of importance, namely the Activity of Interest (AoI). The chapter is divided as follows: the first section presents the *AoI Score* to highlight the limitations of the existing performance metric. The second section introduces the *Ranking Capability Score* (RCS). The third section demonstrates a case study for evaluating the ranking capability of the SWA system whilst the fourth section conducts a quality based evaluation to examine the proposed performance metric against its intended purpose. Finally, this is followed by a comparative evaluation between the (RCS) and $AoIScore$ over three separate scenarios.

Chapter 4 will introduce a new level of assessment. This is to examine the knowledge representation issues for the proposed performance metric. The first section will examine the " Ranking Capability Score" (RCS) versus the three qualitative states which are likely to occur during a real time operation; the first is the *Good State* where the SWA system is ranking the identified activities as perfect as the ground truth. The second is the *Degraded State* where the SWA system is ranking the identified situation as not perfect as the ground truth, and the third is the *Bad state* where the SWA system is ranking the identified events as opposed to the ground truth. The second section introduces an enhanced method for the proposed metrics, called the *Ranking Capability Score'* (RCS'). The third section will conduct a reliability based evaluation, with the help of a case study based scenario, to evaluate the proposed enhanced scoring scheme against the operator perception. The fourth section will present a comparative evaluation and, finally, we will discuss our findings and future work.

The first contribution of this Chapter 5 is in introducing the scheduling capability score (SCS) for evaluating the ranking capability of a SWA system. Furthermore, this work conducts a comparative evaluation with existing performance metrics; the evaluation methodology encompasses two levels of assessments, with the first level being a case study based evaluation. The second level is a quality-based evaluation used to examine the underlying metrics against their intended purpose. The second contribution of this work is to deliver a method to analyse the computational complexities involved in evaluating the prioritisation and scheduling processes for two distinct operations. The first operation is during the assessment stage, where the evaluation process required computes only the necessary values to assess the ranking capability of the real time system. The second operation is during the optimisation stage, where the evaluation process is required to compute more values to assess all ranking instances for any given scenario.

Chapter 6 Concludes the thesis and indicates possible research directions that this work can take.

Chapter 2

Background and Literature Review

2.1 Introduction

The analyst team is required to keep up to date with all five different domains in a dynamically monitored environment, these being classified as sea, land, space, air and cyberspace. To meet the needs of these domains, the data fusion community have introduced the information fusion reference model[27] [26], [35], [55], [66] [72]for multi-disciplinary areas, describing the theoretical concept of a real time system with multiple levels of situational assessment. Each level performs a contextual task during a real time operation. In return, the proposed output of these simultaneous processes is either prioritised in a list of events (namely the tracking activity) or it is represented by visual means, supporting timely responses during real time operation.

This chapter discusses the theoretical concept of a real time system by reviewing a relevant reference model. The next section discusses the literature reviews, highlighting four research problems regarding the capability of the SWA system. The first research problem concerns knowledge based issues, the second concerns ranking capability issues, the third concerns the process refinement stage for assessing the performance of real time systems and the fourth concerns the knowledge representation issue regarding

user perception. As this thesis is focused on evaluating the ranking capability of real time systems, the last section will further discuss the theoretical concept of ranking, prioritisation and scheduling processes from the information fusion perspective. This is to explain how each process is likely to occur in multilevel situational assessments.

2.2 Real Time System

This section discusses four different reference models which have been designed for supporting decision-making resources during a real time operation. Generally, these theoretical models describe the fundamental concept of the construction of real time systems with multiple functions and processes.

The underpinning concept for the SWA system originated from Endsly's SWA reference model. Simultaneously, the joint director of laboratories (JDL) designed a more constructive model for assessing multi disciplinary areas during real time operation, known as the data fusion JDL reference model.

Hence, the data fusion community linked a number of theoretical concepts or standards to create a guidance model with multiple processes for describing the real time system. This section discusses the underpinning concept for these reference models, explaining the essential background of a real time system.

This section is divided as follows. The first section discusses the theoretical concept of the Endsly SWA model, the second section explains the data fusion reference model, the third section describes the SWA reference model, and, finally, the last section describes the decision-making cycle for time sensitive operations.

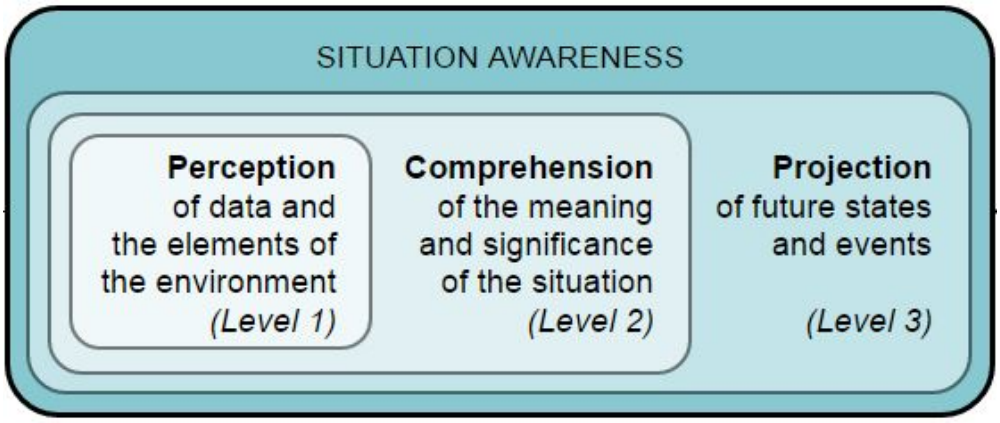


Fig. 2.1 The Endsley (1995) Situation Awareness Model Taken from [27]

2.2.1 Endsly SWA Model

The Endsley's SWA model [27] [26] clearly stands as the reference for real time systems and it is constructed with a three-level mental representation: perception, comprehension, and projection. Endsley has created the core concept of real time systems, stating that the SWA system has the ability to identify, comprehend, and project critical elements of information for multi-disciplinary areas; this is shown in Figure 2.1. Each process performs a contextual task during a real time operation as follows:

- Level 1: Perception of critical element of information.
- Level 2: Comprehend its meaning.
- Level 3: Projects its Future Status.

During the perception stage the real time system is conducting classification, aggregations and correlation techniques to perceive high abstract views for any emerging situation during the real time operation. The comprehension stage further assesses the perceived situation to comprehend its meaning, concerned with the dynamically monitored environment. Finally, the projection stage proceeds beyond the comprehended

situation to anticipate its future developments, hoping to meet the goals of the SWA system it has originally been created for.

This section has discussed the core concept of Endsley SWA system; the next section explains the relationship between the SWA and data fusion information domain.

2.2.2 Joint Director of Laboratories (JDL) Model

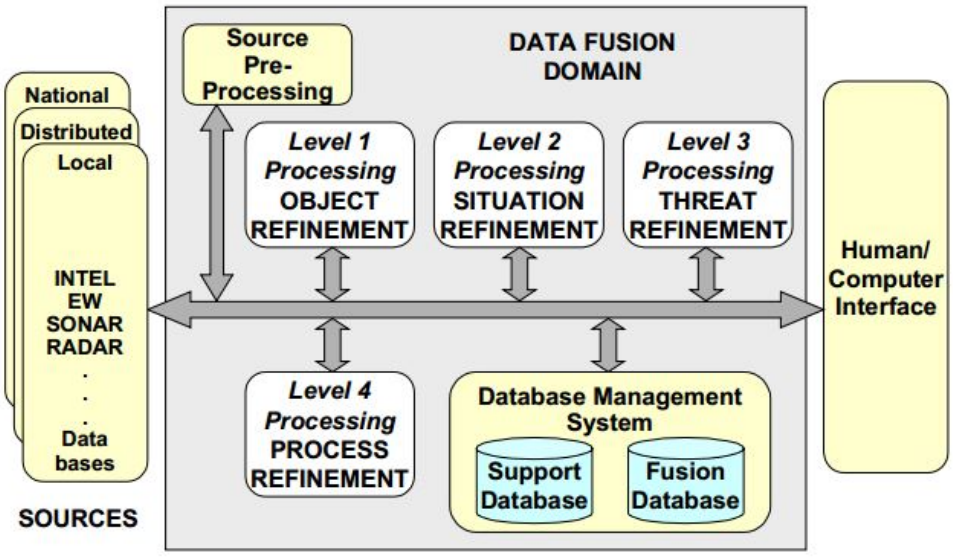


Fig. 2.2 The Joint Directors of Laboratories (JDL) Model Data Fusion (1992 version) Adapted From [55]

The data fusion community [84], [53],[54], [73],[75] [82] [83] [98] [99] [88] has introduced multiple procedures of SWA as shown in Figure 2.2. This is to facilitate an active model for assessing multidisciplinary areas during real time operations. Hence, the researchers over time have combined the core concept of Endsley SWA and the information fusion domain as shown in Figure 2.3.

The first two levels of the data fusion domain are related to the perception stage of the SWA reference model, where the system is receiving the data from multiple sensors to prompt the highest abstract views of the emerging situation. Level 2 and

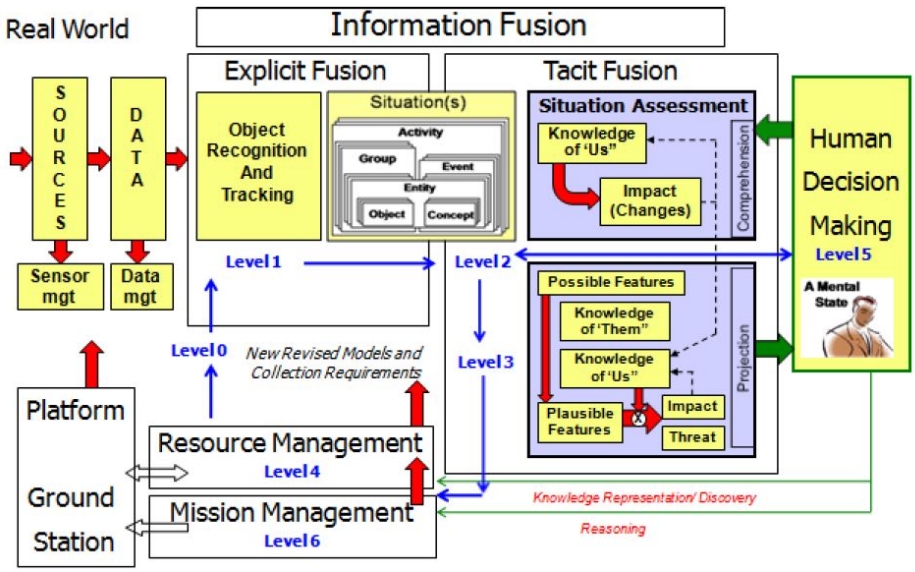


Fig. 2.3 Situation Awareness And The Information Fusion Domain [14]

level 3 are related to the comprehension and projection stages, respectively, and level four is related to the resource management or, in other words, is related to the process refinement stage, where the underlying process is intended to assess the performance of previous levels of the SWA system, that is to assess or improve the data fusion task during a real time operation. Level 5 is concerned with user refinement to address knowledge representation and reasoning for the decision-making resources.

This section has discussed the relationship of Endsley's SWA model and the information fusion domain. The next section discusses the SWA reference model.

2.2.3 Situational Awareness Model

Researchers from the data fusion community have developed a refined concept for the combined version of the Endsely SWA reference model and the data fusion information domain. The enhanced version of the SWA system is shown in Figure 2.4.

The underlying reference model describes the functionality of a real time system, specifically when the SWA system is configured to process multiple sensor information

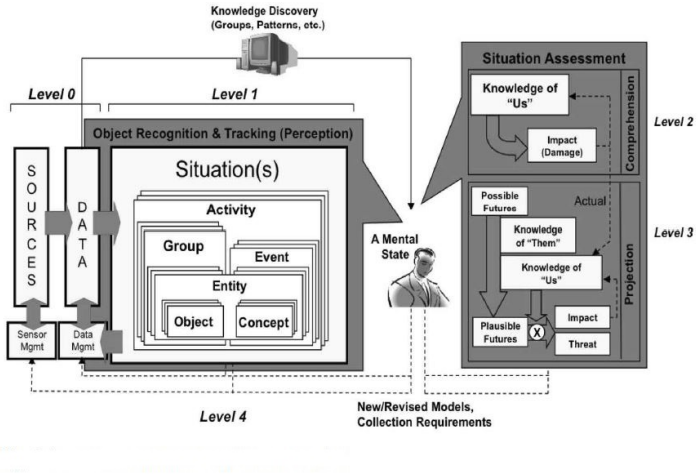


Fig. 2.4 Situation Awareness Reference Model Adapted from [90]

during a real time operation. The perception stage encompasses two levels from the data fusion domains; Level 0, where the underlying process represents the raw sensor information from multiple sources, and level 1, representing the recognition and tracking process. Hence, the perception stage is expected to process multiple raw sensors information for generating a list of tracking activities concerning the identified situations. Level two is expected to comprehend the perceived tracking activities by consulting different pieces of information concerning the dynamically monitored environment. The next level (projection stage) is expected to further process the identified tracking activities by consulting more information to rank them into a contextual order. The fourth level of a real time system is expected to revise all previous stages for supporting the decision-making resources during a real time operation.

This section has discussed the SWA reference model, specifically during the time when the real time system is configured to report different tracking activities during a real time operation, in order to support the decision-making resources. The next section discusses the relationship between the underlying model and the decision-making cycle.

2.2.4 Decision Making Cycle and Time Sensitive Operation

The decision-making cycle of observe, orient, decide and act, has been inspired by a military strategist, as shown in Figure 2.5. The Boyd's [18], [17] decision cycle, or OODA loop, allows the decision-making resources to make appropriate decisions more quickly than one's opponent.

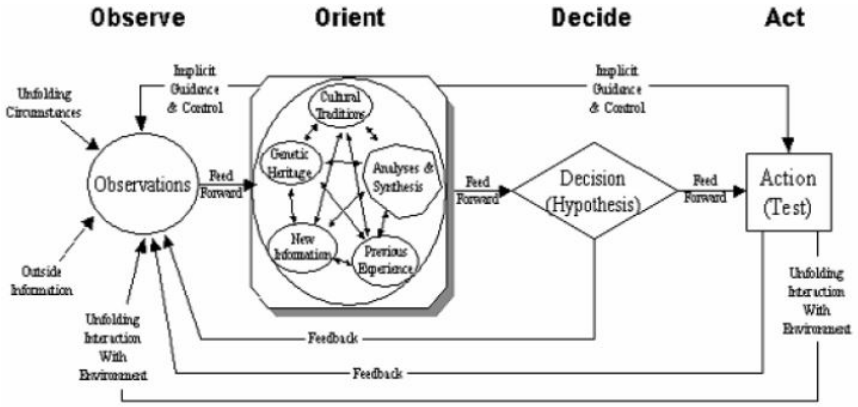


Fig. 2.5 Obsearve, Orient,Decide and Act: The Decison Making Cycle Adapted From[17]

The SWA system is designed to support decision-making resources, for promoting fast responses over a complex situation. With this in mind, research has combined the OODA loop decision cycle with the SWA system, as shown in Figure 2.6.

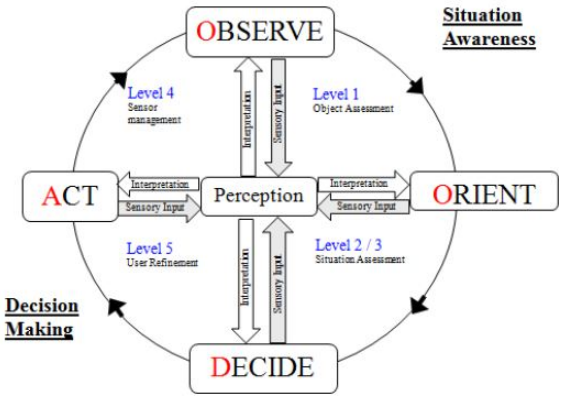


Fig. 2.6 The Decison Making Cycle And Situational Awareness Adapted From[7]

Due to the fact that the SWA system is dealing with a time sensitive operation, the time window for responding is very small, as shown in Figure 2.7.

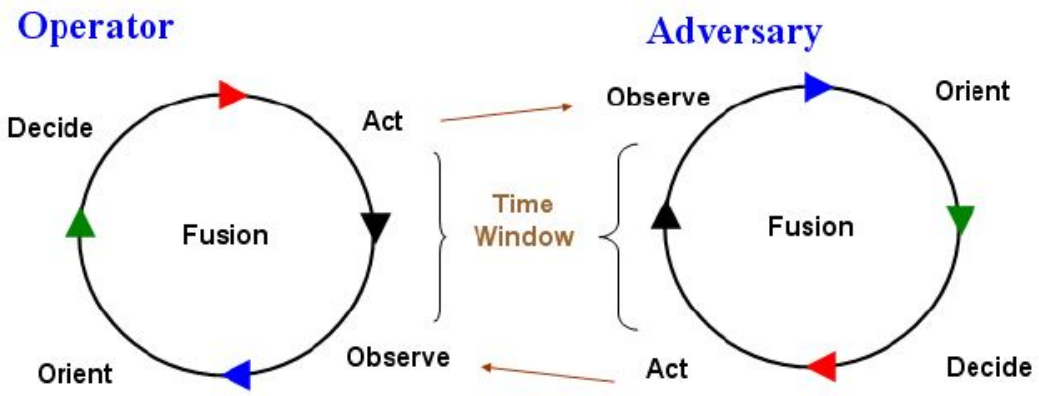


Fig. 2.7 Operator and Adversary OODA Loops With The Associated Time Window To Act[9]

Furthermore, Blach [9] has classified the timely responses paradigm into three main categories: reactive, proactive, and preventive, as shown in Figure 2.8.

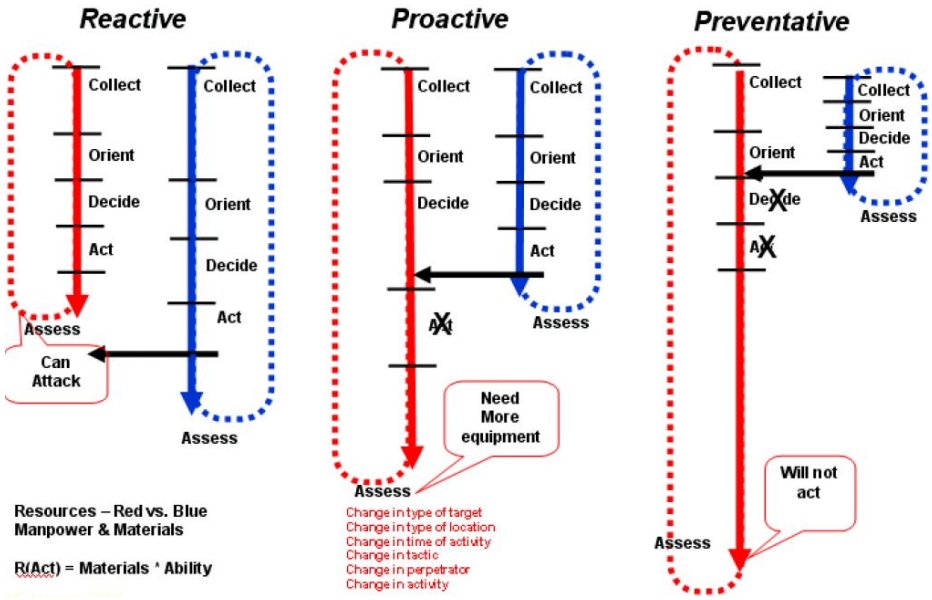


Fig. 2.8 Decsion Making Paradigams In Relation To A Time Sensitive Operation [9]

The reactive response is related to the perception stage, where the contemporary situation has already occurred; in such a case we are detecting. The comprehension stage is related to the current state for the identified situation; in such a case we are responding. Finally, the projection stage is related to the future state, where we are preventing a thing that may or may not happen.

This section has discussed the theoretical concept of a real time system and explained the functionally of each level. In addition to this, it has explained the number of decision-making paradigms concerning the time sensitive operations occurring during a real time operation. The next section explains the related research issues concerning a real time system.

2.3 Capability of Situational Awareness Systems

The previous section described the theoretical concept of a real time system. Furthermore, it showed how the SWA system has been constructed and designed from multiple levels of situational assessments with each level lending itself to assess the next one. With this in mind, there are five different domains using the underlying system; each domain has a different configuration for assessing real time operations. Even within a particular domain, such as the cyberspace area, each SWA system has different needs and configurations in order to assess a contextual situational assessment during real time operation successfully.

Unfortunately, as most of these systems rely on predefined knowledge, they inherit a knowledge based problem and, therefore, their capability in terms of processing new and unexpected information becomes limited.

Hence, the capability issue of an SWA system can lead to other research problems. These include, but are not limited to, the following. The first issue is the reliability of the proposed assessment outputs of the real time system, and the second research

problem is the method of refinement as it may be required for evaluating the capability of the SWA system. The third research problem is the knowledge representation issue for the decision-making resources.

This section discusses four different research problems concerning the capability of real time systems. This section is divided as follows: the first section reviews related work concerning the knowledge based issue of a SWA system. The second section discusses the ranking capability issues concerning the reliability of the proposed assessment output of a real time system. The third section explains the concept of the process refinement stage concerning different assessment methods for evaluating the capability of a real time system. Finally, the fourth section discusses the knowledge representation problem concerning user perception.

2.3.1 Situational Awareness System and Knowledge Based Problem

The knowledge based problem can impact the real time system capability, in terms of processing or assessing a new emerging situation during a real time operation. This is from the lowest level, such as the perception stage, up to the highest level, such as the projection stage. This section reviews an existing SWA system from the cyberspace domain literature [48] [64] [70] [71] , [63], [52] to explain how the predefined configuration of the real time system can impact its capability, specifically when the situational assessment model is assessing a new emerging situation.

This section is divided as follows. The first section discusses the different SWA systems concerned with perception stage configuration. The second section reviews different systems concerning the comprehension stage implementation and, finally, the third section reviews different approaches concerning the projection stage of the SWA system.

Perception Stage

The perception stage processes emerging information from multiple sensors to recognise the high abstract views of any emerging situation during real, or near real, time operations. It does this through applying contextual aggregation and classification techniques to process the multiple data sources from the multiple sensors, as small as atomic levels. Then, it interprets this into a contextual form based on a predefined configuration. Hence, the predefined knowledge can be problematic when dealing with new unexpected information.

This section reviews cyberspace domain literature[29] [56] [96] [22], regarding how the knowledge based issue can impact the capability of the perception stage. Moreover, researchers from the cyberspace domain community are using different components, tools or approaches for assessing real time operations. This section reviews three different and well known approaches currently used for implementing the perception stage. This section is divided as follows. The first part discusses the attack graph and vulnerability approach, the second part discusses the alert correlation approach and, finally, the third part discusses the attack modelling approach.

Attack Graph and Vulnerability

Generally, the attack graph technique and vulnerability scanner is a well known approach in the cyberspace research community. It is mainly used to asses atomic information from multiple sensors, such as Intrusion Detection System, Firewall and Vulnerability Scanners. During the perception stage, the information is interpreted into a contextual attack scenario based on predefined attack templates created for assessing any emerging threat against the dynamically monitored environment. The following examples explain the perception capability problem further.

The author in this work [48] has used attack graph techniques to process the atomic level information from the vulnerability network scanner and then interpreted to a contextual multi stage attack scenario. In other words, the high abstract views of the emerging situation were presented in the form of an attack graph. Interestingly, the knowledge in an attack graph was combined with the network topology information to recognise a possible attack path. Furthermore, the author [64] combined the attack path information with critical asset information to filter out the most important event.

These efforts combined multiple elements within the cyberspace using an attack path, vulnerability scanner information and other elements to perceive any emerging situation during a real time operation. However, the context of data analysis is based on a limited assumption, such as if multiple alerts lay at a predefined attack path, the system will treat those alerts as urgent. Unfortunately, this assumption can be misleading, especially if multiple false alarms are laid on the predefined attack path. The SWA system treats those alerts as urgent while the actual attack itself could take a different path.

Ultimately, predefined knowledge can easily limit the capability of the perception stage during a real time operation. The next section reviews another approach used to support the perception stage for assessing multiple sensors information during a real time operation.

Attack Modelling

The attack modelling approach has been used by the cyberspace research community to facilitate the perception stage in processing multiple sourced information into high abstract views in the form of a contextual attack scenario. It also allows the perception stage to track the intrusion activity from one point to another inside the protected environment. The following examples highlight how such an approach is being impacted

by the knowledge based problem due to the reliance on a predefined knowledge of known attack scenarios.

The author here has implemented the attack modelling approaches to allow the perception stage to track different intrusion activities, thus exploiting different network vulnerabilities. These efforts [47] [52] [70] [71] , [63], [52] have a good vision on predicting the next step of the attacker.

The reference model of tracking the intrusion activity during a real time operation can be misleading to the ground truth of actual attacker behaviour. Advance and complex attack techniques utilise different paths to leverage traditional defence systems. Hence, the attack modelling approach is capable of assessing only known and predefined attack behaviours.

The next section reviews how knowledge based issues can also affect the alert correlation approach when the real time system is processing multiple sensor information to support the perception of the cyberspace situational assessment.

Alert Correlations

Usually when the real time system is receiving atomic level information from multiple sensors, these alerts have no common semantic relationship in order to track the high abstract views of any intrusion activity occurring during a real time operation; most of these alerts are false due to wrong detection.

Researchers [29] [56] [96] [22] have built additional systems to manage IDS alerts based on predefined knowledge of well documented attacks. Other efforts [56][22] have utilised the knowledge of network and topology information to recognise attacks and reduce the number of alerts.

The author [29] proposes alert correlation systems to address the high volume of IDS alerts. The proposed method relies on an unsupervised system (Autonomy). The

system consisted of four components. The first is the normalisation step, to convert heterogeneous alerts into a unified form. Secondly, the date undergoes pre-processing to remove redundant alerts. Thirdly, the attack scenario extraction algorithm is applied to combine knowledge base and its statistical relationship. Finally, the fourth component, the Hidden Markov Model (HMM), does not take into consideration the information from the cyberspace environment, such as network topology, network vulnerability and system configuration.

The prediction paradigms of these efforts [94][24][60] relies on predefined patterns of attack behaviour. Such approaches treat ongoing events in isolation to the enterprise mission, therefore, the data analysis is limited to known attacker behaviour.

The author [22] proposes a threat prediction framework based on alert correlation. The system takes three elements into consideration. First, it uses attack profiles to generate network attack specifics. Secondly, attack plans are used to generate a cover-ability tree. Thirdly, it contains the primitive attack (PA), which is a hierarchical attack class. The first two elements are represented by Coloured Petri nets (CPNs) to assess the security situation assessment and attack scenarios prediction.

The author [96] proposed a prediction framework based on a Dynamic Bayesian Network (DBN) to identify attacker intention and predict future attacks. The proposed framework takes advantage of the Bayesian nature to dispose sequences of data. The system observes the developmental action and varying behaviour at an hourly basis.

The author [56] proposed an alert aggregation system based on the theoretical concepts of an artificial immune system and dangerous theory. During specified time windows, the system opens 6 temporary groups: three groups are for network based IDS alerts and the remaining groups are for host based IDS. If one of the groups reaches the capability to raise the dangers alarm, based on predefined rules, the system only reports those alerts.

Existing efforts [22][96][56] developed correlation engines based on a single source of data analysis (IDS alerts). Such approaches are limited to the performance of network intrusion detection systems. The proposed alert correlations system verifies whether predefined attacks are taking place or not. These systems rely on predefined rules. We found that these rules are either generic or very specific. An intermediate approach is required to avoid short cut heuristic decision making.

The next section reviews different SWA systems to further explain the knowledge based problem concerning the comprehension stage of a real time system.

Comprehension

The comprehension stage is the second level of assessment after the perception stage. During this stage more information resources are consulted to further assess any emerging situations. Furthermore, it applies a contextual perceptive measure to prioritise the identified tracking activities into a predetermined order. This level of assessment comprehends the current state for any emerging situation during real time operations.

Researchers from the cyberspace domain have separated current damages and their anticipated threats [73][74] based on their time of occurrence.

Furthermore, researchers have developed a number of perspective measures to rank the identified tracking activities based on their current state. These include, but are not limited to, the following perspective measures: most serious [58], most important events [87], depth of attacks, breadth [88] and reliability. [86].

The most serious scenario[58] ranks different attack scenarios. The scenario which received the highest number of alerts is classified as most serious. In most important events [87], the attack scenario with less random states received higher scores than others. Depth of attacks [86] involved measuring how close the adversary was to its

possible targets based on the number of hops; the attack scenario with the least number of hops received high scores.

Breadth [88] is used to measure how much of the entire attack scope has taken place in predefined attack templates; attack scenarios with greater scope received higher scores. Reliability [86] is used to measure how sure it is that the attack will take place by taking into account the number of alerts and weight from the past adversary course of action.

ECCARS [58] is a SWA framework used to rank most serious scenarios from the indication of live alert streams. The hierarchical tree of a single attack scenario dynamically ranks attack templates between the value of [0,1]. The attack scenario which received the highest credibility is ranked as most important.

INFERED [86] refine the ranking mechanisms through [58] applying relative entropy (a measure of randomness) to the evolving situation of attack scenarios. Each attack scenario has a universal state with multiple, equally likely, constant states of attacker steps. When one or more of the live alerts matches a predefined attack step, the credibility of the attacker step will increase, thus making the randomness of relevance attack scenario decrease; attack scenarios with less random states received higher ranking.

INFEREDv2 [58] (BN and HMM) Model the stepping stones of attacks in ARENA simulation and dynamically group multiple attackers to single attack templates, called the attack template guide. The author applied relative entropy to rank how likely different attack tracks were to continuously occur on the environment.

The next section discusses the third level of assessment, where the projection stage has been designed to anticipate the future state for any identified tracking activities during a real time operation.

Projection

During the projection stage, the real time system further assesses the emerging situation by applying a contextual perceptive measure to prioritise the identified tracking activities into a predetermined order. This level of assessment anticipates the future state of any emerging situation during real time operations.

Researchers from the cyberspace domain have developed a number of perspective measures based on the future state. This includes the following perspective measures; most likely [41] [30] [31] [25], most plausible

The most likely perspective measure [41] [30] [31] [25] is used to measure the attack's intentions; attack scenarios with greater opportunity and capability to privilege inside the protected network will receive higher intention scores.

In the most plausible measure [42][75], the attack scenario with great opportunity and capability is compared to the intention of attacker scores. The attacker intention is used as a reliable score to solve a conflict when capabilities and opportunities are combined; attack scenarios with high plausibility would receive higher scores.

Most threatening [98] [97] is a measure used to fuse the state of the attacker with the state of the network, to rank different attack scenarios. Most vulnerable [21] is a measure used to project potential exploited vulnerabilities by the attacker; attack scenarios with greater opportunities to exploit the underlying vulnerability will receive higher scores.

TANDI [41] is a threat assessment framework used to inform analysts of possible proactive and preventive action. Capabilities and opportunities are fused to determine the intent of the attacker. TANDI predicts malicious activities one step ahead of the attacker, before the attacker can reach its ultimate goal.

FuSIA [42] is a threat assessment framework in which the capabilities and opportunities are fused to determine the severity of the attacker's next actions. The impact

assessment is combined with the impact estimation, providing information on potential threats. Yang [98] has integrated INFERED (the states of attacker) with TANDI (the state of network) to inform analysts about potential threats.

Salerno et.al[75] have integrated FuSIA with INFERED to refine the ranking of plausible scores. In return, the integrated approach would report future actions two steps ahead of the attacker, before the attacker can achieve its ultimate goal. The plausible scores are based on the assigned manual weight for multiple features such as the capability, opportunity and intent.

F-VLMM [25] is an enhanced approach regarding the attack projection algorithms of the VLMM model.The author has combined the VLMM model with Sugeno Fuzzy logic. The refined approach independently characterises multiple features, such as capabilities and opportunities, which are used by FuSIA [42] to resolve conflicts raised when they are used individually. The result shows good ranking of future attacker progression inside the networks.

Bayer and Yang [21] have extended the VLMM projection algorithm in [30]. The extended approach has implemented training and projection algorithms simultaneously, for an arbitrary number of alerts. In response, the system was able to project potential exploited vulnerability by the attacker.

Unfortunately, we found the existing work to be limited to two views. The first view is when the researcher anticipated future threats based on predefined attacker step and manually assigned weights for ranking the identified tracking activities into a predefined order.

In the second view, the researcher is predicting future intrusion activity based on an analytical solution; this is to allow a wider view to anticipate a greater number of situational assessments by generalising the underlying projection techniques. A

combined approach of the two methods would provide better outcomes for the projection stage.

However, the capability of the real time system is still in question, especially when the SWA system is dealing with new emerging situations or unexpected new information during a real time operation.

This section has reviewed the different approaches and configurations for the three levels of assessment in a real time system; the perception stage, the comprehension stage and finally the projection stage. Furthermore, we have found that the knowledge based issue, or the predefined configuration, can impact the capability of a SWA system. The next section further explains the capability of a real time system in terms of ranking different tracking activities during a real time operation.

2.3.2 Ranking Capability

This section discusses the ranking capability of a real time system with three perspective views; the first view is the linguistic means, second view is the analytic representation and the third view is the operational or practical occurrence during a real time operation.

The scientific concept of ranking, prioritisation, and scheduling demonstrates how these terms occur often in information fusion.

According to the Oxford dictionary[1] the scientific concepts of ranking[1], prioritisation[2], and scheduling[3] have distinct meanings. However, these terms describe overlapping processes and have not yet been appropriately explained for the data fusion community. Figure 2.9 demonstrates how each term occurs in a real time operation.

[1]Position of somebody/something on a scale that shows how good or important they are in relation to other

[2]Prioritize (something) to put tasks, problems, etc. in order of importance, so that you can deal with the most important first

[3]A plan that lists all the work that you have to do and when you must do each thing

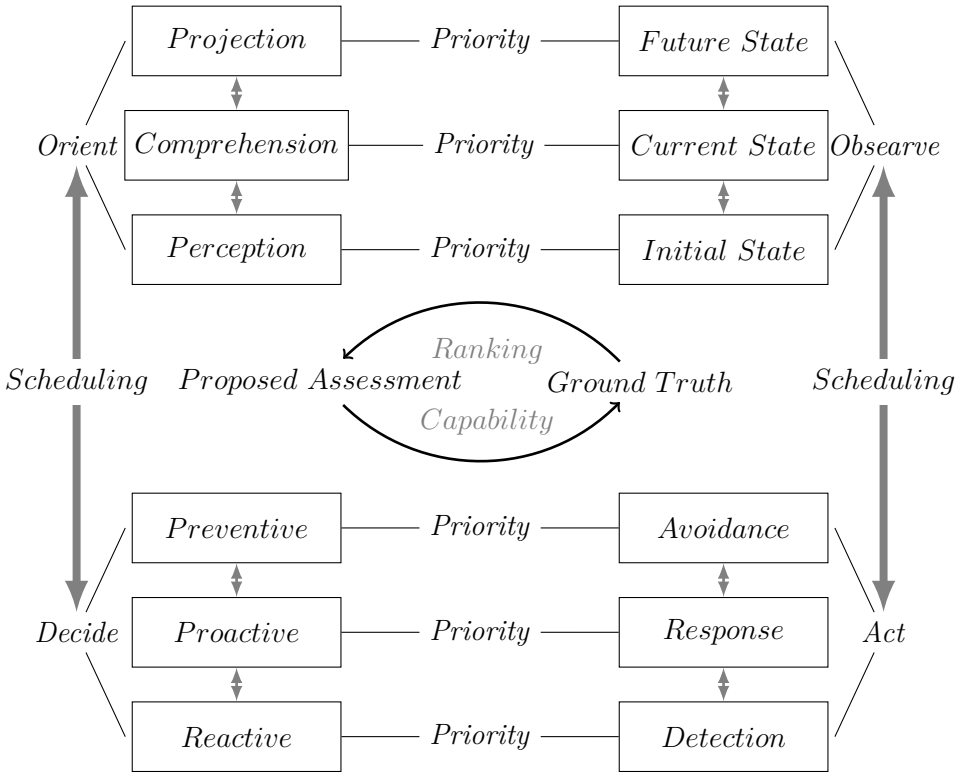


Fig. 2.9 A Reference Model Concerning The Ranking, Prioritisation and Scheduling Processes in a Real Time System

Endsley [27] has defined three abstract views for the SWA system. These levels are shown in Figure 2.9. The first level is the perception of the environment, where the underlying process is prioritising the emerging event based on the initial configuration, classification, and aggregation techniques of the real time system. Simultaneously, the next level is seeking to comprehend the perceived situation; here the system is applying a contextual perspective measure to prioritise the perceived tracking activity based on its current state. Finally, the projection stage seeks to prioritise the comprehended situations based on their future threats.

It is important to mention that when the system is prioritising the emerging situation, the advanced stages are likely to contain the priority list of the previous one. Consequently, the proposed assessment output of a real time system may report more than one priority list during real time operations. Hence, each of these categories should be scheduled based on the available decision-making paradigms.

In a scenario where the analyst team have chosen to conduct a proactive response over the emerging situation, the priority list of the current state should arguably be observed first, before the priority list of the initial stage. Similarly, the priority list for the projection stages should arguably be observed ahead of the comprehension stage. However, there are no definitive answers for which list should be viewed first; the answers depend on many facts, such as the available decision-making paradigms or the interest of the analyst team.

Considering this, researchers [8], [9] from the data fusion community classified the Colonel John Boyd decision-making cycle [18], [17] and the timely sensitive operation into three broad categories: reactive, proactive, and preventive responses. However, during a real time operation, the analyst team may have no option (sometimes) over a complex situation, where the contemporary event has already occurred. In such a case,

we are detecting;the interest of decision making is to perform a reactive response to the detected situation.

Instantly, the comprehension stage is assessing the developments of the identified situation, and while the system is proposing the priority list for the current state, the decision making might be interested in conducting a proactive response over the emerging situation. Simultaneously, the projection stage is processing the emerging situation to predict the future state; in such a case, decision making is important in order to avoid things that will or might happen in the future.

Figure 2.9 shows the occurrence of prioritisation and scheduling processes in a real time system. Ideally, the SWA system is reporting a different priority list concerning a multilevel situational assessment. On the other hand, their desired order to be observed depends on the decision-making paradigms; furthermore, the real-time system may not rank the identified situation appropriately. Section 2.3 discusses a guidance case study during a situation when a real time system is experiencing a ranking capability issue due to a configuration problem in the aggregation process.

Hence, researchers have introduced different performance metrics; this is to assess and improve the ranking capability during a real time operation, as shown in Figure 2.9. Likewise, the assessment process verifies the proposed assessment outputs of the real time system in comparison to the ground truth of the identified situation.

The next section discusses a case study based scenario from existing literature, where a real time system is experiencing a ranking capability issue with the prioritisation and scheduling processes, respectively.

A Case Study

This section illustrates a guidance case study in which the real time system is experiencing ranking capability issues during a multilevel situational assessment. The first

section discusses a case study from the cyberspace domain [74]. The second section demonstrates how the real time system is reporting the identified situation; specifically, when the inline assessment model is reporting fragmented events due to a configuration problem in the aggregation or classification process.

Cyberspace Domains

According to the underlying scenario adapted from [74], the real time system has been configured to represent the emerging situation in the form of a prioritised list of events. However, during the aggregation process of multiple information, the SWA system has identified two classes of events.

Activity of Interest (*AoI*)That is a complex event which has high impact against the protected environment.

Activity That is a complex event which is regarded as a normal *activity*, which has minimal or no impact against the dynamical environment.

Considering this, the real time system has performed two levels of assessment against the identified situation. However, the system is reporting undesirable events due to a configuration problem in the aggregation process. Therefore, the perception and comprehension stages have not ranked the identified events as perfect as the ground truth. The next section demonstrates how the real time system is reporting the identified situation when the system is experiencing a ranking capability issue.

Ranking Capability Issues

Following the contextual scenario in the previous section, the perception stage has reported the identified activities at the time of their occurrence and without any contextual order; the proposed assessment for the perception stage is demonstrated in

Table 2.1. The real time system has conducted a further assessment of the perceived situation. The comprehension stage has ranked the AoI into a contextual order, as demonstrated in Table 2.2.

Proposed Assessment	**Activity**	**Priority**
PA_0	Activity	4
PA_1	Activity	3
PA_2	Activity (AoI)	2
PA_3	Fragmented Activity	-
PA_4	Activity	5
PA_5	Activity (AoI)	1
PA_6	Activity not part of G.T	-
PA_7	Activity	6

G.T Ground Truth

Table 2.1 Proposed Assessment At The Perception Stage Adapted From[74]

It is argued that the ranking instance at the comprehension stage is better than the perception stage, concerning the scheduling process in shifting the AoI over the normal events. Nevertheless, the proposed ranking paradigms are not as perfect as the ground truth, as shown in Table 2.3.

Proposed Assessment	**Activity**	**Priority**
PA_0	Activity	4
PA_1	Activity (AoI)	1
PA_2	Activity (AoI)	2
PA_3	Activity	3
PA_4	Fragmented Activity	-
PA_5	Activity	5
PA_6	Activity not part of G.T	-
PA_7	Activity	6

G.T Ground Truth

Table 2.2 Proposed Assessment at the Comprehension Stage Adapted from[74]

This means that the analyst team will view up to three of the detected activities before looking at the desired class of activity, namely the AoI. However, the decision-

making resources are much closer to the AoI at comprehension stage in comparison to the perception stage.

Ground Truth	**Activity**	**Priority**
GT_0	Activity (AoI)	1
GT_1	Activity (AoI)	2
GT_2	Activity	3
GT_3	Activity	4
GT_4	Activity	5
GT_5	Activity	6
GT_6	Activity	7

Table 2.3 Ground truth for the Identified Situation Adapted from[74]

According to the ground truth (shown in Table 2.3) for the identified situation, the two levels of assessment have not proposed appropriate ranking paradigms for the emerging situation. This, of course, can impact the timely responses during a real time operation. However, going back to the scientific definitions of ranking, prioritising and scheduling terms in section 2.3, we can easily understand each process with the help of the underlying scenario. To begin with, the real time system has identified two classes of complex events, where each of them encompassed prioritised list events. Hence, the ground truth for the identified situation contained two priority lists: the first one is for the AoI, and the second one is for the normal activity; this is shown in Table 2.3.

Taking this into consideration, the action of prioritising different events into an order of importance is scientifically referred to by the prioritisation process, so that the decision-making resources can deal with the most important event first. However, the process of shifting one class of event over another, considering all the lists of events within the desired classes and regardless of their respective degree of importance, is referred to as the scheduling process.

It is important to distinguish between the two terms of scheduling and prioritisation due to the fact they resemble different tasks during a real time operation. When the

multilevel situational assessment does not represent the emerging situation appropriately, the verification process is required to conduct two separate operations to evaluate each process mathematically.

Generally speaking, to evaluate the prioritisation and scheduling capability of the real-time system, the process refinement stage will require comparison of the proposed assessment of each task with the absolute truth of the situation. The action of questioning the capability of a real time system on a scale that shows how good they are in relation to the ground truth is scientifically referred to as the linguistic definition of ranking, and that is provided in section 2.3. Hence, the ranking capability of the real time system is either referred to as the scheduling process or the prioritisation task in information fusion.

This section has explained the three terms of ranking, prioritising and scheduling processes of a real time system. The next section discusses the process refinement stage, which is level 4 of a SWA system.

2.3.3 Process Refinement

The process refinement of the Joint Directors of Laboratories (JDL) is a meta-process used to assess and improve the data fusion task for supporting the decision making resources during a real time operation. In this section, we discuss different approaches designed to assess the performance of a real time system. The assessment process during the process refinement stage (level 4) has been designed to verify the capability of previous data fusion levels. Hence, the verification technique can take two forms of evaluation; qualitative and quantitative assessment.

During the qualitative stage, researchers[19][20] [50][7][11][7][77] have developed a number of methods to investigate the capability of a real time system.

Llinas [54] has linked three types of threatening situations with the analysts responses. The first situation is the reactive one to alert users to immediate threats; the analysts select the appropriate response in a second. The second situation is the proactive situation to detect anomalous behaviour and alert the user to anticipated threats; the analysts select the appropriate response from a few seconds to a minute. The third situation is preventive to prevent potential threats before they reach deployment; the response can take from a few minutes to an hour.

Rebovich [68] explained that operators are often under extreme pressure to make decisions as quickly as possible where there is much uncertainty; this may lead to adopt short cuts heuristic approach. Therefore, the system should support the decision maker paradigms and operators projection needs, in order to facilitate operator goals.

FAIR [19] prioritise the selection of decision maker goals based on the time available for different situational assessments; 1) no time limits correspond to passive action with unlimited goals, 2) tactical situation with several goals of action 3) opportunistic situations with two goals of actions 4) scrambled situation with only one goal available to analyse.

Bayers [21] and Tadda [90] advised that SWA systems should take into consideration the time elapsed before a decision could be made or for action to be taken. For example, in the cyberspace domain, the time window of an attackers next step and its relation to analysts should be investigated.

Blasch [9] evaluates the reliability of a fusion system to deliver a set of information over an appropriate time window. He utilized the queuing theory to evaluate the input of a fusion system until the decision maker becomes available; the aim of the system is to hold the excess data.

C-OODA [20][13] addressed the time for analysts against the level of certainty in a situation; if the level of certainty is high with no time available, then the process of

reasoning over such a situation will be terminated, otherwise, it will keep processing the information in order to reduce the level of certainty.

Blasch [7] used control theory to simulate the iteration of user's effort to reduce uncertainty versus the time available to respond.

Recognition Primed Decision (RPD) model [50] utilised past experience against recognised situations from information fusion. The decision maker goal cues to information system. This results in systems providing reduced time for decision making and promote accuracy.

OODA-RR [36] discussed the interaction between the adversarial course of action and analysts response. Grant presented a list of attacker stages against the corresponding responses of analysts including passive, reactive, proactive and preventive actions.

During the quantitative stage, researchers [15][4][90][76][77] have developed a number of methods to investigate the capability of a real time system.

Salerno[74] proposed a scoring scheme to evalute the capability of SWA systems, in terms of ranking important events into a contextual order. Tadda [89] explained that the further the AoI is from the top of the list, the more time the analysts will need to assess such situations. The scoring techniques provide an indication of how close the analysts are to the most important activity. Unfortunately, the AOIScores are limited to contextual attack scenarios, Blasch [15] suggested for an extension to have wider views for the AoI scores.

Less attention has been given to performance evaluation in regards to the ranking capability of a real time system. Originally, existing performance metrics had not been designed for measuring the ranking capability of the SWA system. Specifically, corner cases of different situational assessments needs and configurations have not been

considered. This thesis presents advanced research work carried out to evaluate the ranking capability of an SWA system for a number of different scenarios.

This section has explained the researched issues concerning the process refinement stage (level 4). The next section explains the difference between the ranking capability concerning prioritisation and scheduling processes of a real time system, where each process uses different mathematical operations to obtain the number of ranking instances for any given scenario.

2.4 Analytical Analysis: Prioritisation and Scheduling

Returning to the case study demonstrated in section 2.3, the ground truth has identified two priority lists, each one encompassing some prioritised events. It is desirable for the event classed first to be observed by the decision making resources before the second one.

To assess the ranking capability of the demonstrated case study, the evaluation process seeks to perform two distinct operations to quantify the ranking capability of the real time system. The first level of assessment is to measure how the system ranks each priority list compared to the ground truth. The second level measures how the proposed assessment output shifts the desired class of event over the other normal one.

This section is divided as follows. The first section discusses the preliminary steps of quality-based evaluation concerning prioritisation/scheduling processes, and the second section introduces two different scoring schemes intended to evaluate the ranking capability of the real time system. The third section conducts a comparison evaluation to examine two distinct performance metrics against their intended purpose.

2.4.1 Number of Ranking Instances

Generally speaking, the notion of different ranking instances concerning the prioritisation process are related to the act of rearranging, or permuting, all the identified events into a sequence or order.

The number of permutations for each priority list can be defined using the operation of factorial ($N!$), where N represents the number of events for each priority list, and factorial ! provides the number of possible ranking instances for those activities. The factorial, usually written as $n!$, denotes the product of all positive integers less than or equal to n. This allows us to define the number of ranking instances concerning the priority list for each class of event. Furthermore, to compute the total number of ranking instances concerning the priority list of the AoI, we can apply the factorial operation $(2!)$ = two ranking instances. Moreover, to compute the total number of ranking instances concerning the priority list of a normal *Activity*, we can also apply the same operation$(4!) = 24$ ranking instances.

However, according to the demonstrated scenario, the perception stage shown in Table 2.1 has reported undesirable events (fragmented activities). Consequently, the total number of events is greater than the two priority lists found in the ground truth shown in Table 2.3. Hence, to compute the number of ranking instances for each priority list we must consider other events being proposed by the real time system. The evaluation process requires us to perform the permutation operation ${}^{n}P_{k} = \frac{n!}{(n-k)!}$, where n is the total number of distinct events being identified by the real time system, and the k is the number of events on the priority list for each type of activity.

Likewise, to compute the total number of ranking instances concerning the priority list of the AoI, we can apply the permutation operation $({}^{8}P_{2} = \frac{8!}{(8-2)!})$. Moreover, to compute the total number of ranking instances concerning the priority list of a normal *Activity*, we can apply the same operation $({}^{8}P_{4} = \frac{8!}{(8-4)!})$.

In fact, the number of the ranking instances using the factorial operation is different than the permutation one; this is because the permutation formula is the product of factorial (k!) and the combination operation ${}^{n}C_{k} = \frac{n!}{k!(n-k)!}$. It is here that the research becomes more interesting. This is because we can recall the relevant ranking instances concerning the prioritisation process or the scheduling process separately. With this in mind, there are three likely situations: in the first scenario, when the decision-making resources are not interested in the order of each priority list, we can only use the combination operation; that is to obtain all the ranking instances concerning the scheduling process.

In the second scenario, when the analysts team is interested in the order of each priority list and their position/time to be observed over other classes of event, we can use the ${}^{n}P_{k} = \frac{n!}{(n-k)!}$ operation. That will give us a greater ranking of instances in comparison to the previous scenarios. In the third scenario, when the prioritisation process is interested only in the order of each priority list, mathematically we can obtain the relevant ranking instance using the factorial operation ($N!$).

This section has explained the difference between the prioritisation and scheduling processes in terms of the analytical means, where each process used different mathematical operations to obtain the number of ranking instances for any given scenario.

2.5 Chapter Summary

This chapter has discussed different reference models to explain the essential background concerning the theoretical concept of a real time system.

The second section has explained three different research problems concerning the capability of a SWA system. The first research problem was the knowledge based issues

and the second was the ranking capability issues. The third issue was the process refinement stage for assessing the performance of a real time system.

Furthermore, as this thesis is focused on evaluating the ranking capability of a real time system, we explained the theoretical concept of ranking, prioritising and scheduling processes in a real time system. It has also explained how each term may occur during a real time operation. Moreover, this work has conducted an analytical analysis to describe further how the prioritisation and scheduling tasks are different, not only by the linguistic mean, but they are also practically distinct in the way that each term can be evaluated mathematically.

Chapter 3

Ranking Capability Score (RCS)

3.1 Introduction

In a dynamical monitored environment, a team of analysts need timely and accurate information in order to respond proactively to complex situations. Typically, there are thousands of reported activities in real time operations. To direct the analyst's attention to the most important activity researchers [22] [44] [43] [49] have performed multiple processes of situational awareness (SWA) to rank the most important activity into a predetermined order. Eventually, the output of the SWA system[4] [61] should be able to facilitate the decision-maker in responding to complex situations efficiently.

According to the SWA reference model there are multiple levels of situational assessments. Each level lends itself to assess the next level simultaneously. The first level is the perception of the emerging situation where the SWA system is processing the detected data from heterogeneous sensors using classification, aggregation and correlation techniques. The proceeding levels further assess the underlying situation, using prioritisation techniques in order to rank the emerging threats into a predetermined order.

Unfortunately, various factors can impact the situational assessment capabilities. One factor is the up to date knowledge about the protected environment due to the dynamical changes of topology information, node connectivity, enterprise mission, network services, critical asset, context of risk assessments model and new waves of contemporary attacks. In real time operations the SWA system often has limited information about the dynamical environment, therefore, the missing information renders the situational assessment incomplete. Consequently, the team of analysts are periodically in the loop of reviewing, updating and optimising the SWA system.

To combat this, researchers have introduced different performance metrics to verify the capability of the SWA system. However, less attention has been given to performance evaluation regarding the prioritisation process. Specifically, the existing performance metric, "*The Activity of Interest Scores*"[74] [15], has not considered corner cases for different situational assessment needs and configurations. Originally, it had not been designed for evaluating the capability of the SWA system in relation to the prioritisation process.

This chapter presents a new performance metric for evaluating the ranking capability of the SWA system, presenting two contributions. The first develops a modelling scheme for representing the outputs of a SWA system in the form of a list of prioritised events. The second contribution introduces the "Ranking Capability Score" (RCS) as well as a guidance case study for evaluating the ranking capability of an SWA system. This will primarily deal with the prioritisation process of a real time system, under a contextual scenario where the SWA system has identified only a number of tracking activities regarded as important, but each with different degree of importance, namely the Activity of Interest (AoI).

The chapter is divided as follows: the first section presents the *AoI Score* to highlight the limitations of the existing performance metric. The second section

introduces the *Ranking Capability Score* (*RCS*). The third section demonstrates a case study for evaluating the ranking capability of the SWA system whilst the fourth section conducts a quality based evaluation to examine the proposed performance metric against its intended purpose. Finally, this is followed by a comparative evaluation between the (*RCS*) and *AoIScore* over three separate scenarios.

3.2 The Activity of Interest Score *AoIScore*

This section introduces the activity of interest score, beginning with an overview of the intended purpose of the *AoIScore*, as well as the contextual scenario it was designed for. Alongside this, an illustrative case study will be demonstrated to explain the evaluation process for the prioritisation technique. Next, corner cases of the original scenario will be demonstrated before quantifying the ranking capability using the *AoIScore*. Finally, a comparative discussion focused on the results between the two scenarios highlights the limitations of the underlying metric.

3.2.1 Overviews

The SWA system at the detection stage (Level 0) receives heterogeneous information from multiple sources; however, not all detected events are relevant to the current situation. The job of the perception stage (level 1) is to differentiate between the activities of interest and irrelevant activities.

The problem here is that the system might report activities of interest as they appear without giving any contextual order. Therefore, the task of the comprehension stage (level 2) is to perform further assessment of the identified situation. Additionally, after conducting a contextual risk assessment through consulting different pieces of

information within the dynamical environments, the SWA system ranks the AoIs into a predetermined order.

Arguably, each level provides better understanding for the emerging situations, maximising the benefit for the decision-maker. The $AoIscores$ have recently been used to quantify the ranking capabilities for each level of an SWA system. In particular, it measures how well different levels are ranking the activities of interest over the least important ones.

Researchers [66], [67], [74], [15] have adapted a systematic method to quantify the ranking capabilities of an SWA system. The first step is to quantify the ranking capabilities for the perception stage and then for the comprehension stage. The obtained scores are intended to quantify the amended assessments for each stage respectively to provide another dimensional support for the decision-maker. It provides an insight into how the systems perform in ranking important activities into a predetermined order for each level of the SWA system.

The performance metric, namely the AoIs $scores$, requires two components in order to function. These are the absolute truth for the emerging situation, called *ground truth*, and the proposed ranking paradigms of SWA system in real time operation, called *proposed assessment*.

The next section introduces a case study for which the underlying performance metric was originally designed.

3.2.2 Case Study 1: SWA System has Identified Mixed of Cyber Activities

In this section we demonstrate the original scenario which has been adapted by [74]. The underpinning situational assessment has the following properties:

1. The situational assessment has been adapted from the cyberspace domain.

2. The underlying SWA system has identified two types of Cyber activities at the perception stage (level 1);

   (a) The first is a complex event which has high impact against the protected environment, namely $The\ Activity\ of\ Interest$.

   (b) The second is a complex event which is regarded as a normal activity, which has minimal or no impact against the dynamical environment.

3. After conducting further assessment at the comprehension stage (level 2), the underlying SWA system did not rank the activity of interest as perfect as the ground truth.

4. The $AoI\ Scors$ was used to measure how well each level ranked the AoIs over normal activities.

The next section demonstrates an illustrative example for measuring the capability of the underlying scenario.

3.2.3 Measuring the Ranking Capability of an SWA System

Following the contextual scenario in the previous section, the perception stage reported the identified activities as they appeared and without any contextual order. The proposed assessment is shown in Table 3.1. The comprehension stage ranked the activities of interest into a contextual order; the results are seen in Table 3.2.

Apparently, the ranking instance at the comprehension stage is better than at the perception stage, in terms of shifting the $AoIs$ over the normal events. Nevertheless, the underlying ranking paradigms are not as perfect as the ground truth, as shown in Table 3.3.

After introducing the components for the underlying situational assessments, the next step evaluates the ranking capability for each level of an SWA system, using the *AoI Score* [74]. This is defined in equation 3.1:

$$\text{AoI Score} = \frac{\text{NAoI} * \text{NA} - \sum_{i=1}^{NAIR} p_i}{\text{NAoI} * \text{NA} - \sum_{i=1}^{NAoI} i} \tag{3.1}$$

where

$NAoI$ = Number of AoIs in Ground Truth

$NAoIR$ = Number of AoIs in Results

NA = Number of Activities in Ground truth

Pi = Position of the i^{th} Activity of interest

The *AoI Score* is intended to quantify the performance of an SWA system through the ranking of the activity of interest. It has been used a number of times [66] [67] [15] within the data fusion community.

Table 3.1 Proposed assessment at the perception stage adapted from[74]

Proposed Assessment	**Activity**	**Priority**
PA_0	Activity	4
PA_1	Activity	3
PA_2	Activity (AoI)	2
PA_3	Fragmented Activity	-
PA_4	Activity	5
PA_5	Activity (AoI)	1
PA_6	Activity not part of G.T	-
PA_7	Activity	6

G.T Ground Truth

After applying the *AoI Score* (defined in 3.1) to the proposed assessment for the perception stage, the SWA scores (0.33). This means the administrator will view two thirds (2/3) of the detected activities before looking to the important one.

Table 3.2 Proposed assessment at the comprehension stage adapted from[74]

Proposed Assessment	**Activity**	**Priority**
PA_0	Activity	4
PA_1	Activity (AoI)	2
PA_2	Activity (AoI)	1
PA_3	Activity	3
PA_4	Fragmented Activity	-
PA_5	Activity	5
PA_6	Activity not part of G.T	-
PA_7	Activity	6

G.T Ground Truth

Then again, after applying *AoI Score* (defined in 3.1) to the proposed assessment at the comprehension stage, the SWA system scores (0.78).

Table 3.3 Identified activity at the ground truth adapted from[74]

Ground Truth	**Activity**	**Priority**
GT_0	Activity (AoI)	1
GT_1	Activity (AoI)	2
GT_2	Activity	3
GT_3	Activity	4
GT_4	Activity	5
GT_5	Activity	6
GT_6	Activity	7

This means the administrator is much closer to the activities of interest. Thus, the SWA system at comprehension stage is much closer to the ground truth (shown in Table 3.3), in comparison to the perception stage.

The underlying performance metrics have successfully quantified the extended scenario, as shown in Figure (3.1), specifically, under a scenario where the situational

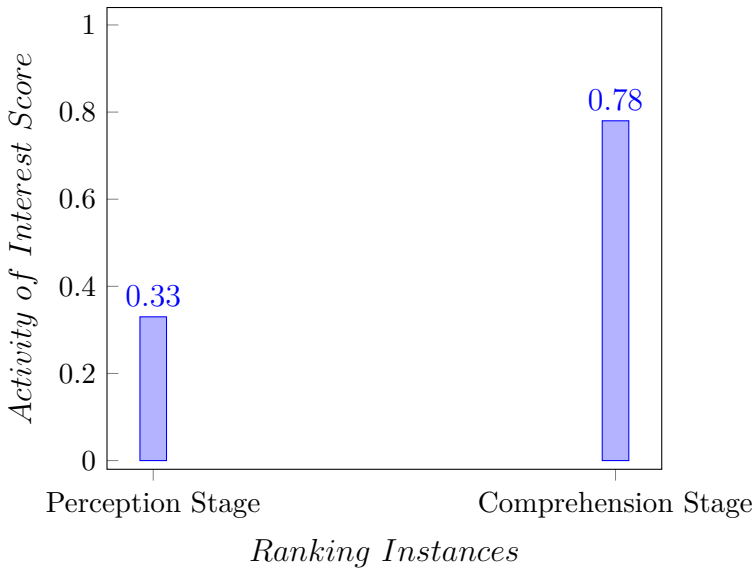


Fig. 3.1 Case Study 1: Quantifying The Ranking Capability of SWA System Using The *AoIScore*

assessment has proposed two classes of activities. Furthermore, the perception and comprehension stages scored 0.33 and 0.78, respectively; this can be seen in Tables 3.1 and 3.2. Evidently, the ranking of the comprehension stage is better than the perception stage.

The following section examines the underlying performance metric under an extended scenario, where the SWA system has been configured to report only the AoIs but with different severity paradigms.

3.2.4 Case Study 2: SWA System has Identified Only Important Activities

This section does not change the original situational assessment (demonstrated in previous section). However, this section is considering corner cases from the original scenario where the SWA system in real time operation may identify only the AoIs, but with different severity paradigms. Furthermore, we consider an emerging situation where the SWA systems might be instructed to report only the AoIs.

Consequently, the proposed assessment at the perception stage will contain only the AoIs. Typically, the underlying process for the identification stage reports the AoIs as they appear and without contextual order. The task of the comprehension stage is to conduct further assessment in order to prioritise the most important activities into a predetermined order. Ultimately, the analysts under this scenario will view the most important activity first, followed by the least important ones respectively. Hence, this chapter examines how well the *AoI Score* will quantify different ranking instances for the underlying situational assessment.

To facilitate us in performing the evaluation operation using *AoI Score*, we have randomly proposed two different ranking instances to a set of three AoIs. The first ranking instance will represent the proposed assessment at the perception stages (shown in Table 3.1), while the second ranking instances will represent the proposed assessment at the comprehension stage (shown in Table 3.2) and, finally, the absolute truth for the extended scenario will be represented (shown in Table 3.3).

3.2.5 Measuring the Capability of Situational Assessment

This section will evaluate the capability of the *AoI Score* in evaluating corner cases of different situational assessments.

We will apply the *AoI Score* algorithm to evaluate the proposed assessment for the perception stage (level 1) and the comprehension stage (level 2). Finally, we will compare the two scores to quantify the amended assessment for each level of the SWA system.

The evaluation process begins by quantifying the ranking capability for the perception stage. First, it will extract relevant values from the proposed assessment as shown in Table 3.5. Secondly, it will substitute relevant values for the ground truth, as shown

Table 3.4 Pre-determined order for identified Activity at Ground Truth

Ground Truth	**Activity**	**Priority**
GT_0	Activity (AoI)	1
GT_1	Activity (AoI)	2
GT_2	Activity (AoI	3

in Table 3.4. In order to satisfy the parameters of "$AoI\ Score$" defined in 3.1, we will substitute the first two values from the ground truth $NAoI = 3$ and $NA = 3$.

Table 3.5 Proposed assessment at the perception stage(Level 1)

Proposed Assessment	**Activity**	**Priority**
PA_0	Activity (AoI)	3
PA_1	Activity (AoI)	2
PA_2	Activity (AoI)	1

Next we obtained two remaining values from the proposed assessment; these are the position of the i^{tt} activity of interest $P_i = 6$ and the geometric sum of AoI $_i = 6$. Finally, we quantify the capability of the situational assessment for the perception stage using the AoI score as follows: $AoIscores = \frac{(NAoI*NA)-P_i}{(NAoI*NA)-i} = \frac{(3*3)-6}{(3*3)-6} = \frac{3}{3}$ or 1.

After computing the ranking capability for the perception stage we quantify the ranking capability for the comprehension stage. Likewise, the evaluation process will extract four values from the ground truth, (shown in Table 3.4) and the proposed assessment for the comprehension stage (shown in Table 3.6).

Table 3.6 Proposed Assessment at the comprehension stage(Level 2)

Proposed Assessment	**Activity**	**Priority**
PA_0	Activity (AoI)	2
PA_1	Activity (AoI)	1
PA_2	Activity (AoI)	3

Because the ground truth remains the same for the underlying situation, we extract only relevant values from the proposed assessment for the comprehension stage. The first value is i^{th} position of the AoIs which is $P_i = 6$. The second value is the geometric sum of AoIs, $i = 6$. Finally, we compute the capability of situational assessment against the comprehension stage as follows: $AoIscores = \frac{(NAoI*NA)-P_i}{(NAoI*NA)-i} = \frac{(3*3)-6}{(3*3)-6} = \frac{3}{3}$ or (1).

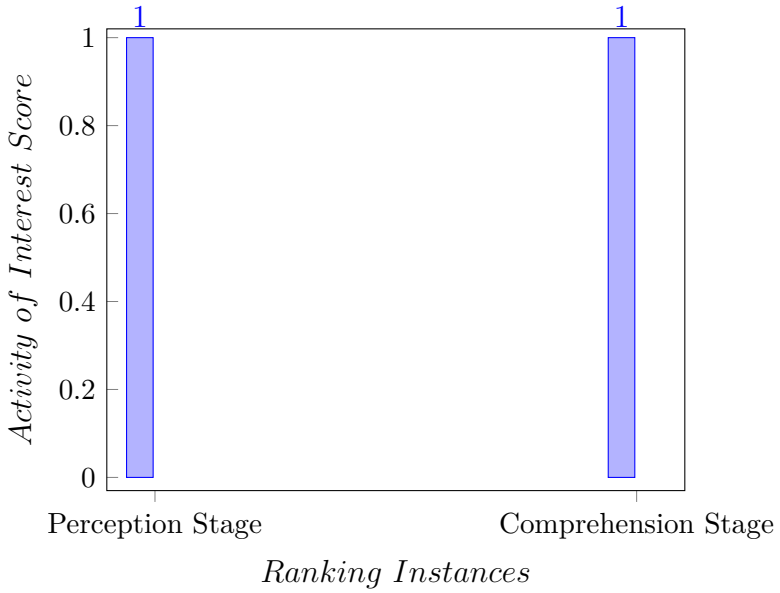


Fig. 3.2 Case study 2: Quantifying The Ranking Capability of SWA using the *AoIScore*

After we have applied *AoI Score* to the ranking instances for the perception stage,(provided in Table 3.5) we get the underlying metrics scores 1. Furthermore, we applied it again to the output at the comprehension stage (as shown in Table 3.6); the algorithm also scored 1.

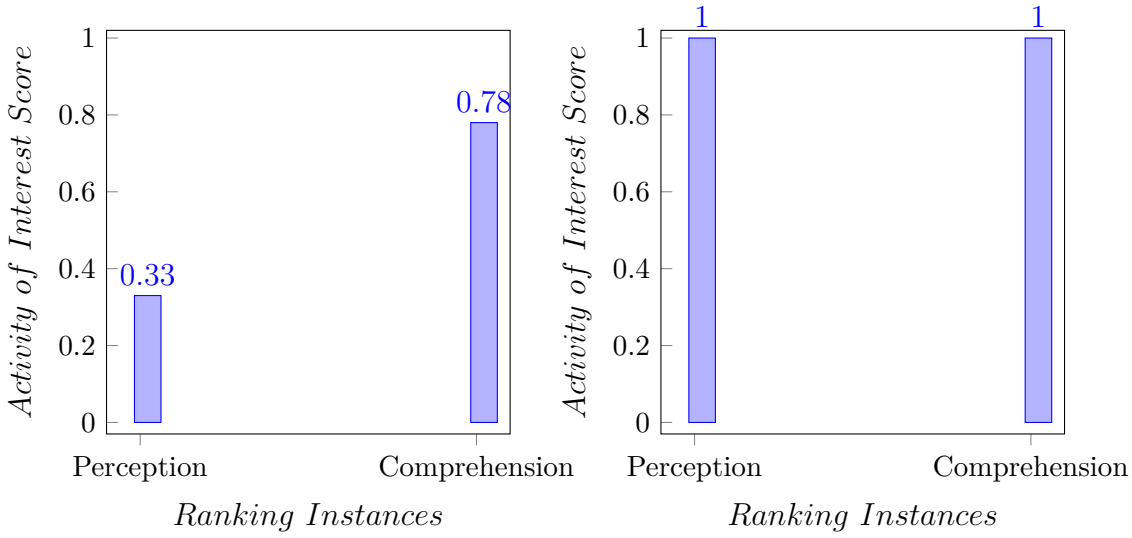
Although the ranking instances at the perception and comprehension stages are different, the *AoI Score* provides similar scores for both stages, indicating that both levels of the SWA have ranked the activity of interest as perfect as the ground truth (as shown in Table 3.4).

Unfortunately, the underlying metrics do not quantify the amended assessment for different levels of the SWA system. Furthermore, our initial evaluation for the

AoI Score poses some threat to common belief, in which the underlying metrics have provided inappropriate scoring schemes for multiple levels of situational assessments (shown in Figure 3.2).

The next section presents a comparative result based on the case study based evaluation for the *AoIScore*.

3.2.6 Comparative Results



(a) Case Study 1:The SWA System has identified mixed of activities

(b) Case Study 2:The SWA system has identified only important activities

Fig. 3.3 Evaluating the *Activity of Interest Score* under an extended scenario

We examined the performance metric, namely the *AoIScore*, under an extended scenario where the SWA system identified only important activities but with different severity paradigms. Unfortunately, the results showed that the underlying performance metric does not quantify the ranking capability for different levels of the SWA system. Rather, it treats all the identified threats only as important activities, regardless of their respective severity paradigms. This is because the underlying metric was originally designed to evaluate a contextual situational assessment. The SWA system

has identified two classes of a threatening situation. The first is the AoI, with significant impact against us, while the second is normal activities with minimal or no impact against the dynamical environment.

In Figure 3.3 (Case Study 1) part (A) the $AoIScore$ has quantified the proposed assessment at the perception stage (provided in Table 3.1) with the value of (0.33). This is indicating the SWA system has ranked the identified threats as not perfect as the ground truth (shown in Table 3.3). Moreover, it has quantified the proposed assessment for the comprehension stage (shown in Table 3.2)with value of (0.79), indicating that the SWA system has not ranked the identified threats as perfect as the ground truth (shown in Table 3.3). Apparently the ranking instances at the comprehension stage are better than the perception stage. Hence, the underlying metric has successfully quantified the underlying situational assessment.

However, in Figure 3.3 (Case Study 2) part (B) the underlying metric has quantified the proposed assessment at the perception stage (provided in Table 3.5) with value of (1), indicating that the SWA system has ranked the activity of interest as perfect as the ground truth (given in Table 3.4). Similarly, it has quantified the proposed assessment at the comprehension stage (provided in Table 3.6) with value of (1) indicating that the SWA system has ranked the identified threats as perfect as the ground truth (given in Table 3.4). The result in Figure 3.3 part (B) showed that the AoI score is not able to quantify the amended assessment for each level of SWA system, specifically under an extended scenario where the SWA system has identified only important activities but with different severity paradigms. In regards to this, the underlying performance metric has treated all the identified threats only as important activities, regardless of their respective severity paradigms. Hence, it does not quantify the amended assessment performed at multiple levels of an SWA system.

In summary, the *AoI Scores* is not designed for situational assessment where the SWA system is instructed to report only the AoI, or in a situation where the underlying system has identified only an important activity but with different severity paradigms.

Due to these circumstances, the next section introduces *The Ranking Capability Score*. The underlying metric is intended to evaluate the ranking capability of an SWA system, specifically under a scenario where the SWA system has identified only the AoIs but with different degrees of importance.

3.3 Ranking Capabilities of SWA System

In this section we provide an overview for the prioritisation process of the SWA system. First, we explain the intended purpose of the performance metrics, namely "*The Ranking Capability Score*", as well as the contextual scenario it was originally designed for. The second part introduces the modelling scheme for the extended scenario. Finally, the third part will discuss the development phase of the underlying performance metrics.

3.3.1 Overviews

The SWA system has different layers of complexities. Therefore, an advanced SWA system is one which encompasses different levels of situational assessments. Each level lends itself to rank the most important activities into a predetermined order to maximise the benefit for the decision-makers. The first level of an SWA system is the perception stage. At this stage the system is reporting the identified activities of interest at the time of their occurrence and without contextual order. The next level, which is the comprehension stage, introduces ranking into a predetermined order.

Originally, the system seeks to rank the activity of interest based on their severity paradigms.

However, there are different factors which can impact the ranking capabilities of the SWA system. For that reason, we have developed performance metrics to evaluate the ranking capability of different SWA systems. This, in particular, is for a scenario where the SWA system is configured to report only the AoIs, or for a situational assessment where the SWA system has identified a finite number of activities but with different severity paradigms.

In general there are five different domains [93] [95] [9] [51] where people are required to keep up-to-date with dynamically changing environments such as Air, Sea, Land, Space and Cyberspace. It might be reasonable to assume that distinct domains might have different responses to deal with threat situations. However, all these domains share a similar aim, which is to maximise the benefits for the decision-makers.

Despite that, even within one specific domain, such as cyberspace, researchers [28], [86], [31], [42], [62] have different designs, configurations and purposes for their systems. Furthermore, each system may also have different perceptions about what is an AoI. With this in mind, we have decided to generalise the modelling of the situational assessment, in the hope that researchers, [6], [38] from various disciplines can adapt the proposed performance metrics to their individual needs regardless of variation of different system configurations, or, views on what are deemed important activities.

Therefore, the next section will model the situational assessment scenario, in which the SWA system has identified only a finite number of AoIs, but each with different degrees of importance.

3.3.2 Modelling the Situational Assessment

Generally speaking, the situational assessment encompasses two main components. The first is the absolute truth about the situational assessment, namely the *ground truth* and the second is the *proposed assessment* by the underlying system.

In this section we model the situational assessment components concerned with the prioritisation process of the SWA system. Specifically, we represent mathematically the outputs of the SWA system under a scenario where the underlying system has identified a finite number of events but with different severity paradigms. First of all, we define the relevant objects for the *ground truth*, and then we discuss how the *proposed assessment* might represent those elements.

From the data fusion perspective, usually, the team of analysts have a number of complex events, each one representing the highly abstract views of an element within the monitoring environment. Different events are likely to have distinct priorities. With this in mind, the absolute truth about the situational assessment is defined in equation 3.2:

$$GroundTruth = \{AoI_i\} \tag{3.2}$$

where $1 \leq i \leq N$

TheAoI_i have significant impacts, initially, it had a different priority and this is defined in equation 3.3:

$$AoI_i \in \{1, 2.....N\} \tag{3.3}$$

where $1 \leq i \leq N$

The AoIs represent a set of complex events while the i^{th} represents the respective priority of each one. Therefore, the first activity AoI_1 found on the ground truth is

prioritised with the value of (1), then the next immediate activity is prioritised with the value of (2) and the least important activities found on the ground truth is prioritised with the value of (N), wherein, N = total number of events being proposed by the situational assessment.

As we are modelling a situational assessment concerned with the prioritisation process, the ranking of the AoIs with regard to their severity paradigms is crucial to the underlying scenario. For that reason, we introduce the *number of hops*, NoH_i. Each complex event will have a predetermined order and this is defined in equation 3.4:

$$NoH_i = N - i \tag{3.4}$$

where, the total number of hop $ToH = N - 1$ and $i = 1, ..., N$

The NoH_i indicates how many hops away from the bottom of the list each $i'th$ activity is. Conversely, it will show the predetermined order for each AoI with respect to their initial priority. For example, when considering a number of complex events which may occur in the ground truth list, the least important activity should remain on the bottom of the list with 0 number of hops. The next immediate important activity should have a number of hops equal to 1. Hence, the number of hops is incrementing by one up until the most important activities. The most important activity of all should have the maximum number of hops, which we call the ToH, and this is defined above in equation 3.4 as $ToH = N - 1$.

Ideally, during the situational assessment in real time operation, the most important activity which is found on the ground truth should have the highest score in terms of importance, hence the score of importance is related to the severity of cyber events.

Therefore, we have introduced the score of importance SoI_j for each complex event as found on the ground truth and it is defined in equation 3.5:

$$SoI_j = \{N - NoH_i\}, SoI_1 > SoI_2 > .... > SoI_N \tag{3.5}$$

where, $1 \leq j \leq N$, $1 \leq i \leq N$

According to the underlying situational assessments we have defined three objects for the ground truth. The first object is the AoI_i, containing a finite number of distinct complex events, each one of which represents the high abstract views of a particular element within the protected environment. The second object, NoH_i, represents the predetermined order for each complex event, and the third objective, SoI_j, represents the score of importance which is relevant to the severity paradigms of the complex event. Therefore, the absolute truth of the situational assessment for the underlying scenario is defined below as:

$$GroundTruth = \{AoI_i \subseteq SoI_j, NoH_i\} \tag{3.6}$$

where $1 \leq i \leq N \quad AoI_i \rightarrow SoI_i$

After we have defined the absolute truth for the situational assessment, next, we need to define how different SWA systems might propose the identified activities during situational assessment. As previously explained, we do not know how different systems will rank the AoIs but there are three likely situations: the proposed assessment might rank the AoIs as perfect as the ground truth, or, it might rank them as not perfect as the ground truth, or, in the worst scenario, the system might rank the AoIs as

opposed to the ground truth. Therefore, due to the presence of uncertainty, we define the ranking for AoIs as a random set. This is defined below as:

$$AoI_r = \{AoI_1, AoI_2.....AoI_N\} \tag{3.7}$$

where the set $r \in AoI_i \;\; r|1 \leq r \leq N$

In fact, if the SWA system has proposed different rankings for the AoIs, in comparison to the ground truth, then the related score of importance will also change accordingly. In regards to this, we have defined the SoI_j as a set and it is defined below as:

$$SoI_r \in \{SoI_1, SoI_2.....SoI_N\} \tag{3.8}$$

Where the set $r \in SoI_j \;\; r|1 \leq r \leq N$

It is equally important to mention that the $AoI_r \ni SoI_r$. Hence, whilst the SWA system is experiencing issues with ranking capabilities, the underlying system may not rank the AoI_r as perfect as the ground truth. Therefore, if the ranking paradigms for the AoIs have been changed, then the SoI_r will also change accordingly.

Finally, the last object previously defined is the NoH_i, and it is likely to remain as the same as the ground truth, as long as the number of important activities remains the same. Therefore, the *ProposedAssessment* PA for the underlying situation is defined below as:

$$ProposedAssessment = \{AoI_r \subseteq SoI_r, NoH_i\} \tag{3.9}$$

where $1 \leq r \leq N, \;\; AoI_r \ni SoI_r, \;\; 1 \leq i \leq N$, the set $r \in AoI_i, \;\; r \in SoI_j$

We have introduced a new modeling scheme as a way of encouraging researchers from various disciplines to develop mathematically and to share various contextual situational assessment models to address different needs. The underlying concept behind the modeling scheme is to overcome a number of challenges [83] [10] in the data fusion literature as follows:

Challenges at the researcher level [87], [42], [88]; different researchers neither share their risk assessment model information nor provide their real situational assessment outcome. However, they may provide some generic information about their progress in assessing different situations.

Challenges at the open source level[45]; SWA systems existing commercially are not configured by default to perform a multilevel assessment from the data fusion perspective. Furthermore, some open source SWA systems [92] [91] [69], [16], claimed to have the ability to identify threat situations at the perception level, and some of these systems [2, 3] may also claim to perform some features at the comprehension level. However, to enable these systems to actually perform multilevel situational assessments requires the development of necessary features for both the perception stage (level 1) and comprehension stage (level 2).

Unfortunately, such implementation encompasses the development of classification, aggregation and correlation techniques for the perception stage. In addition to that, they require the development of a contextual risk assessment model for a dynamic environment, as well as a contextual perspective measure at the comprehension stage.

Regrettably, after all these developments, these systems still do not provide the necessary components for the proposed performance metric, such as the *ground truth* and *proposed assessment* outcomes.

Challenges at enterprise level [59]; most enterprises are not willing to share their own risk assessment model information or their situational assessment outcomes. For

that reason, the proposed modelling scheme was the solution aimed at overcoming these barriers and encouraging researchers [37] to develop an abstract view for different situational assessments concerned with a high level of data fusion.

We are aware that some sectors[23] are offering some SWA system services on a small scale, such as at the enterprise level, or on a big scale, such as at national level. Usually, these sectors rely on the process refinement stage; mathematical representations of the dynamic environment. Our modelling approach will help these sectors to apply the proposed performance metrics during the process refinement stage for evaluating different SWA systems at any scale.

After introducing the modelling scheme for the extended situational assessment, with regard to the prioritisation process, the next section presents the development phase for "*The Ranking Capability Score*". Two unsuccessful attempts during the development process are also presented.

3.3.3 Performance Metric

This section defines the relevant parameters for the performance metric "*The Ranking Capability Score*", as well as another two unsuccessful attempts during the development process.

According to situational assessment for the underlying scenarios, each AoI has a predetermined order based on their severity paradigms. Though, to measure the ranking capabilities for the SWA system, first we need to determine the actual number of hops AoH_i from the proposed assessment outputs, and this is defined below as:

$$AoH_i = \{ToH - NoH_i\}, AoH_1 > AoH_2 > .... > AoH_N \tag{3.10}$$

where, $1 \leq i \leq N \quad 1 \leq h \leq N$

The above equation will obtain the actual number of hops for each i^{th} activity being proposed by the SWA system. Next we obtain the current number of hops CoH_i for each complex event as it was proposed by the SWA system; this is defined below as:

$$CoH_i = N - AoI_r \tag{3.11}$$

The set $r \in Aoi_i$ And $r|1 \leq r \leq N$

Once we have obtained the CoH_i and AoH_i, we will subtract the values of AoH_i from CoH_i. In return, we will get the difference in number of hops for each i^{th} activity, which we call DoH_j. This is defined below as:

$$DoH_j = \sum_{i=1}^{j}(AoH_i) - \sum_{i=1}^{j}(CoH_i) \tag{3.12}$$

where, $1 \leq i \leq N \quad 1 \leq j \leq N$

Interestingly, the DoH_j will give us the current ranking state based on the pre-determined order of each i^{th} activity as it was proposed by the SWA system. Yet, to compute the ranking capabilities for the underlying situation, we need to obtain the reference number of hops, RoH_i, for each activity as it was proposed by the SWA system. This is defined below as:

$$RoH_i = SoI_j * DoH_j \tag{3.13}$$

where, $1 \leq i \leq N \quad 1 \leq j \leq N$

The next step is to determine the overall reference number of hops for the identified list of activities. We call it the *Ranking Capability Score RCS*. It is defined below as:

$$RCS = \sum_{j=1}^{N} \left(\sum_{i=1}^{j} (SoI_i)\, DoH_i \right) \tag{3.14}$$

where, $1 \leq i \leq N \quad 1 \leq j \leq N$

This part has introduced the development phase for the underlying performance metric *RCS*. The next section examines the proposed performance metrics against its intended purpose, with the help of case study 2 demonstrated in section (3.2). The case study is under an extended scenario where the SWA has identified only important activities in a real time operation but with different severity paradigms.

3.4 Measuring the Capability of Situational Assessments

According to the extended scenario in section(3.2), the situational assessment has identified only the AoI on the ground truth, as shown in Table 3.4.

Initially, we substitute relevant values for the *ground truth* before determining relevant values for the *proposed assessment* for both the perception and comprehension stage. We quantify the ranking of the SWA system at the perception stage, where the AoI appear and without contextual order (shown in Table 3.5) and at the comprehension stage where the AoI are ranked into a contextual order (shown in Table 3.6.) Finally, we compare results between the two levels.

Ground Truth

We extract the ground truth objects as shown in equation number 3.6. The first object is the AoI_i and it is defined in equation 3.3, while the second object is SoI_j as shown in equation 3.5, and the third object is NoH_i and is defined in equation 3.4.

We have substituted relevant values from the ground truth output as shown in Table 3.4. The first object is the activity of interest $AoI_i = \{AoI_1(1), AoI_2(2), AoI_3(3)\}$. Where $N = 3$. The second object is the score of importance $SoI_j = \{soi_1(3), soi_2(2), soi_3(1)\}$ and the third object is the predetermined order of each complex event $NoH_i = \{noh_1(2), noh_2(1), noh_3(0)\}$, where $ToH = 2$.

Proposed Assessment

We extract the proposed assessment objects shown in equation number 3.9. The first object is the proposed ranking for the activity of interest AoI_r and it is defined in equation 3.7. The second object is the score of importance SoI_r and it is defined in equation 3.8 and the third object is NoH_i as defined in equation 3.4.

Perception Stage Level 1

We have substituted relevant values from the proposed assessment at the perception stage as shown in Table 3.5. The first object is the activity of interest $AoI_r = \{AoI_r(3), AoI_r(2), AoI_r(1)\}$, where $N = 3$. The second object is the score of importance $SoI_r = \{soi_1(1), soi_2(2), soi_3(3)\}$, and the third object is the predetermined order of each complex event $NoH_i = \{noh_1(2), noh_2(1), noh_3(0)\}$, where $ToH = 2$.

Comprehension Stage Level 2

We have substituted relevant values from the proposed assessment at the comprehension stage as shown in Table 3.6. The first object is $AoI_r = \{AoI_1(2), AoI_2(1), AoI_3(3)\}$. The

second object is the score of importance $SoI_r = \{soi_1(2), soi_2(3), soi_3(1)\}$. The third object is the number of hops, as shown in equation 3.4 $NoH_h = \{noh_1(2), noh_2(1), noh_3(0)\}$.

Because the $AoI_r \ni SoI_r$, then, when the AoI_r have different ranking instances in comparison to the ground truth, the relevant score of importance SoI_r also moves accordingly. After we have defined the necessary components, such as the *ground truth*, and *proposed assessment* for both the perception stage and the comprehension stage, we can move to the next step, where we quantify the ranking capability for each level respectively.

Quantifying the Ranking Capability for the Perception Stage and the Comprehension Stage

Now, we measure the capability of the situational assessment at the perception stage. To do this, we compute the current number of hops as shown in equation 3.11 $CoH_i = \{coh_1(3-3=0), coh_2(3-2=1), coh_3(3-1=2)\}$.

Secondly, we compute the actual number of hops, as shown in equation 3.10 $AoH_i = \{aoh_1(2-0=2), aoh_2(2-1=1), aoh_3(2-2=0)\}$.

Thirdly, we compute the difference in number of hops, as shown in equation 3.12, $DoH_j = \{doh_1(2), doh_2(2), doh_3(0)\}$.

Fourthly, we compute the reference number of hops for each AoI, as shown in equation 3.13 $RoH_i = \{roh_1(3*2=6), roh_2(2*2=4), roh_3(1*0=0)\}$.

Finally, we compute *Ranking Capability Score* at the perception stage as follows: $(6)+(10)+(10) = 26$, and the *Ranking Capability Score* at the comprehension stage as follows: $(3)+(3)+(3) = 9$

Based on the above scenario, the proposed assessment at the comprehension stage scores better than at the perception stage; the SWA awareness at the perception stage scores (26). This means the underlying situational assessment at this level has ranked

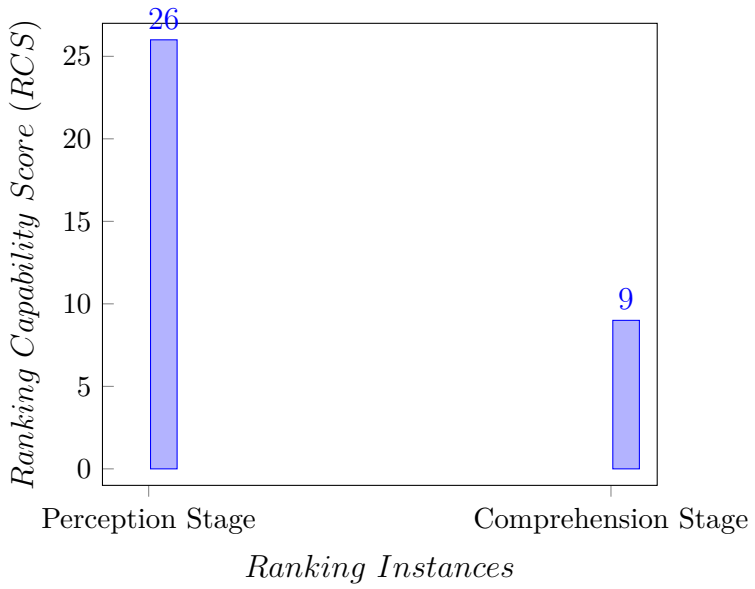


Fig. 3.4 Case Study 2: Quantifying The Ranking Capability of SWA System using the *RCS*

the AoI as opposed to the ground truth. This is because the underlying process at the perception stage is reporting the AoIs based on their time of occurrence.

While this is the case, the proposed assessment at the comprehension stage scores (9). This means the SWA has ranked the AoIs as not as perfect as the ground truth. Perhaps, this is due to capability issues with the system ranking algorithm, configuration or poor risk assessment. Hence, the ranking paradigm at the comprehension stage was better than at the perception stage. Arguably, the underlying metrics has quantified the amended assessment at each level of the SWA system appropriately.

The next section will examine the underlying performance metrics "*RCS*" against its intended purpose. Likewise the evaluation process conduct a quality based evaluation for validating the proposed performance metrics over three separates scenarios. The underpinning evaluation will encompass three phases. The first phase will use an analytical approach to compute the number of ranking instances for any given scenarios concerning the prioritisation process of a real time system. The second phase will use Matlab to simulate all the ranking instances computed during the analytical stage.

The third phase intends to examine the potential of the proposed scoring scheme in terms of providing a unique score for all possible ranking instances proposed by the simulation phase.

3.5 Quality Based Evaluations

This section discusses the quality based evaluations. To begin with, we discuss some fundamental factors to the underlying process.

The first section discusses the characteristics of the SWA system's outputs. In addition to that, it will demonstrate how the number of states are likely to occur for different scenarios that can be obtained.

From the data fusion perspective, the second section reviews different SWA system configurations for the purpose of finding a determination point for the underpinning evaluation.

The third section presents a comparative evaluation for the *Ranking Capability Score* (*RCS*) and the *Activity of Interest Score* (*AoIScore*). over three separate scenarios.

3.5.1 Number of Ranking Instances Versus the Number of *AoI*s

There are different means and ways for maximising the benefits to the decision-maker. For instance, different SWA systems [85] may produce different forms of outputs.

Generally, there are two forms of outputs, the first one being the visualisations techniques [24], [32], [48], where the underlying system produces a number of images, pictures, signals or other visual means to draw the attention of the decision-maker.

The other type of output is traditional, where the system reports lists of activities [15], to draw the analyst's attention to threatening situations. Our performance metric has been designed carefully to evaluate any SWA system and has been configured to report a list of activities for the decision-maker, specifically for the purpose of measuring the prioritisation process capability of different SWA systems.

The notion of different ranking instances is related to the act of rearranging, or permuting, all the AoIs into some sequence or order. For example, if the SWA system has detected three AoIs, there are six permutations for the three distinct activities, which are as follows: AoI_1(A),AoI_2(B),AoI_3(C), namely: 1-(A,B,C); 2-(A,C,B); 3-(B,A,C);4-(B,C,A);5-(C,A,B) and 6-(C,B,A).

Hence, the number of permutations of the AoIs is n factorial (usually written as $n!$), which means the product of all positive integers is less than or equal to n. This allows us to define the number of ranking instances per single scenario using the operation of factorial ($N!$), where N represents the number of AoIs, and the factorial ! provides the number of all possible ranking instances for those activities.

In return, we can obtain the number of all possible states likely to occur for a single scenario. For example, if the SWA system has $AoI = 2$, the total number of state is obtained by $(2!) = 2$ then we will have two ranking instances. Likewise, if there are $AoI = 3$, then the number of ranking instances can be obtained by factorial $(3!) = 6$ ranking instances.

It is apparent when the AoI increases, the number of relevant states increases accordingly. It can be concluded that the potential performance metric should be examined against all ranking instances for a number of scenarios. In the next section we will discuss the dimensional views for the AoI in the hope of defining a determination point for underpinning evaluations.

3.5.2 Determination Point for the Maximum Number of AoIs

In order to set a determination point for the scalability of the proposed performance metrics we must define the contextual concept of *AoI*s.

Complementary to that, we review different SWA systems with the objective of establishing a good understanding of different configurations, thus enabling us to define the maximum number of AoIs per scenario.

We investigated the AoI from both quantitative and qualitative perspectives; qualitative being the contextual concept of AoI and the quantitative being the number of important activities for each qualitative view.

In data fusion literature [10] [33] [54] [82] there is no standard definition for an AoI. However, the underpinning concept for the AoIs is an important activity of great significance. With this in mind, we have found three dimensional views for the AoIs.

The first view is referred to as a situation. Some researchers [50], [19], [13], [36] have classified heterogeneous information from multiple sensors into different classes of situations where each one is directly linked to the decision-maker's response, timeliness and retrospective situation.

Furthermore, each single situation can form a significant risk against us. One single situation is equivalent to a single activity of interest. Quantitatively, existing efforts have defined approximately 1 - 4 different threatening situations, therefore there is a maximum of four AoIs which are likely to occur for this dimensional view.

The second view is an attack scenario. From the data fusion perspective, researchers [25], [87], [98], [97], [75], [41] have modelled the attacker's course of actions into a multi-stage attack. Each attack scenario can form a significant impact, therefore a single attack scenario is regarded as one AoI, hence, the number of AoIs are dependent on the number of attack scenarios which occur during the situational assessments.

Alternatively, again based on the existing literature, researchers [58], [57] have only designed 2-4 attack scenarios for confronting an emerging situation.

The third view is referred to as attacker steps [46],[5] [40]. During the situational assessment for multi-stage attacks, researchers have designed different SWA systems to process heterogeneous information from multiple sources, so as to predict the highest abstract views of attacker steps. In other words, the underlying system will decode different attacker steps into an attacker track.

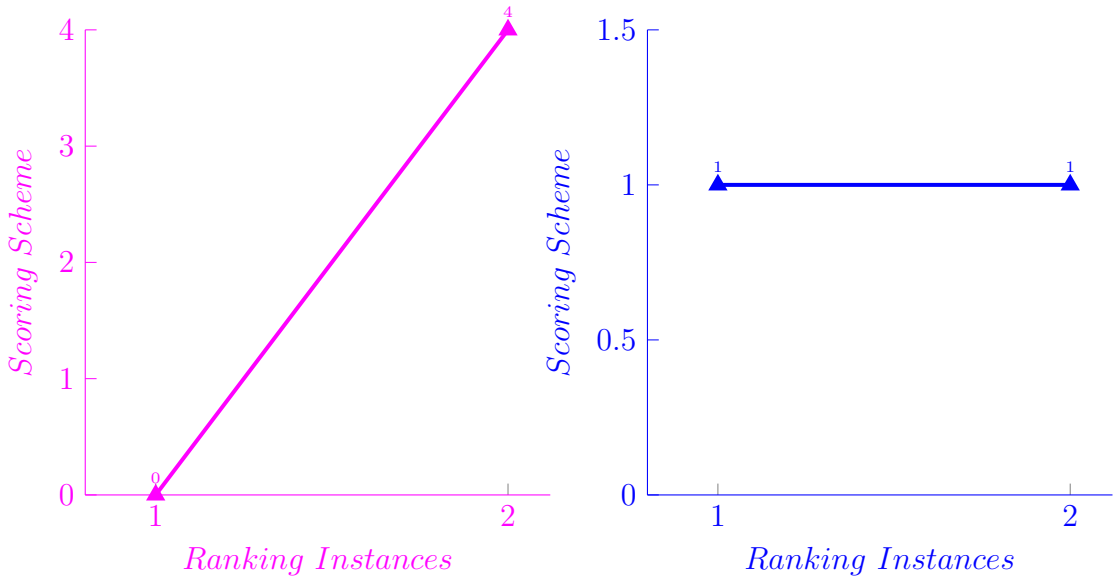
We have found approximately 8 steps for each track; however, during the situational assessment of an emerging situation, researchers found only 3 to 4 steps which might pose significant risk on the targeted environments. Therefore, the maximum number of AoIs at this dimensional level are 2-4 AoIs.

After reviewing the above literature, we found a determination point where the performance metric should successfully scale multiple scenarios with a maximum number of four tracking activities during real time operation. In other wording the proposed performance metric should be able to evaluate the ranking capability for different size of priority lists, specifically where the situational awareness SWA system is reporting a priority list with at least four tracking activity during real-time operation.

The next section presents a comparative evaluation for the *Ranking Capability Score* (*RCS*) and the *Activity of Interest Score* (*AoIScore*). Originally, both performance metrics designed to serve different situational assessments configured to report 2 classes of activities and only important activities, respectively. The evaluation process examines both performance metrics under the second situational assessment in which the SWA system is configured to report only the AoI but with different severity paradigms.

3.5.3 Comparative Evaluation

This section presents a comparative evaluation for the ($AoIScore$) and (RCS). Originally, researchers used the ($AoIScore$) to evaluate the ranking capability of the SWA system, particularly under a scenario where the SWA system has identified two classes of events. The first class contains tracking activity with high impact on the dynamical environment, whereas the second class contains normal activities with low or minimum impact against the protected environment. With this in mind, this section conducts



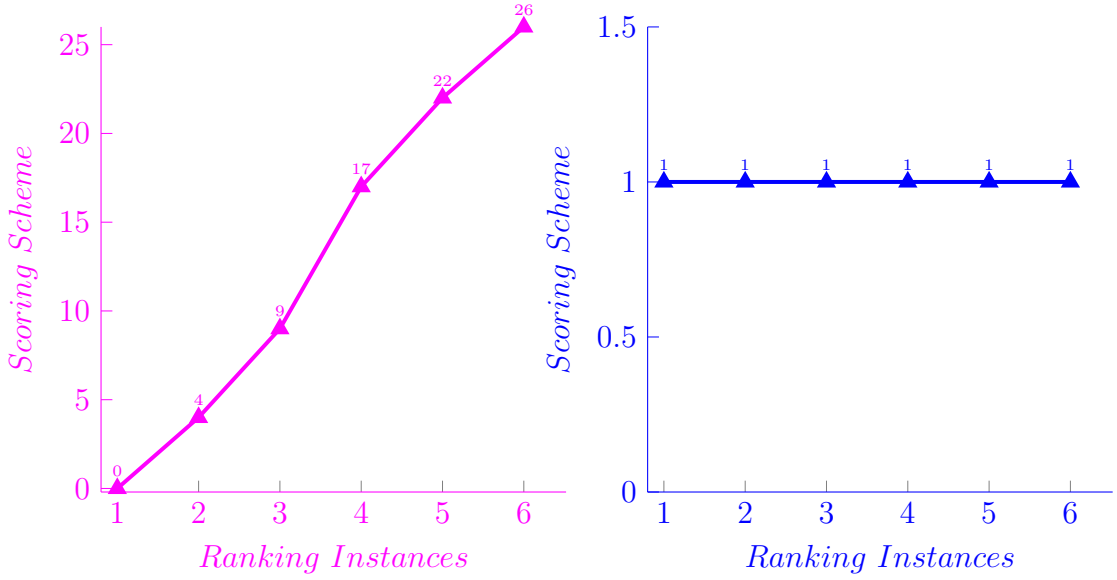
(a) *Ranking Capability Score* (b) *Activity of Interest Score*

Fig. 3.5 Quality based evaluation for RCS versus $AoIScore$. $(2!) = 2$ *Ranking Instances*

a quality based evaluation for the $AoIScore$ under an extended scenario, where the SWA is reporting only the tracking activity of highest interest for the analysts team but with different severity paradigms.

We examine the scoring scheme against three different scenarios where the SWA system reports 2 (Figure 3.5), 3 (Figure 3.6)and 4 (Figure3.7) tracking activity, respectively.

The *AoIScore* in Figures 3.5, 3.6 and 3.7 parts (*b*) (*d*), score (1) for all the ranking instances in all three scenarios. In the first scenario 3.5 the *AoIScore* scores [1] for two different ranking instances, indicating that ranking instances (1) and (2) have ranked the identified activities as perfect as the ground truth. Apparently, the first ranking instances are different than the last. In the second scenario 3.6 the *AoIScore* is [1] for six different ranking instances indicating the emerging threats are ranked as perfect as the ground truth. While, in the third scenario 3.7) the *AoIScore* is [1] for 24 different ranking instances showing that the underlying performance metric doesn't provide appropriate scoring for the the underlying scenario. This is because it was not designed to measure the underlying situational assessment in which the SWA system has identified only important activities but with different severity paradigms. Regrettably, the *AoIScore* treats all the AoIs as only important activities regardless of their respective priority paradigms. Hence, the *AoIScore* does not quantify the underlying situational assessments adequately.



(a) *Ranking Capability Score* (b) *Activity of Interest Score*

Fig. 3.6 Quality Based Evaluation for the *RCS* versus *AoIScore*. (3!) = 6 *Ranking Instances*

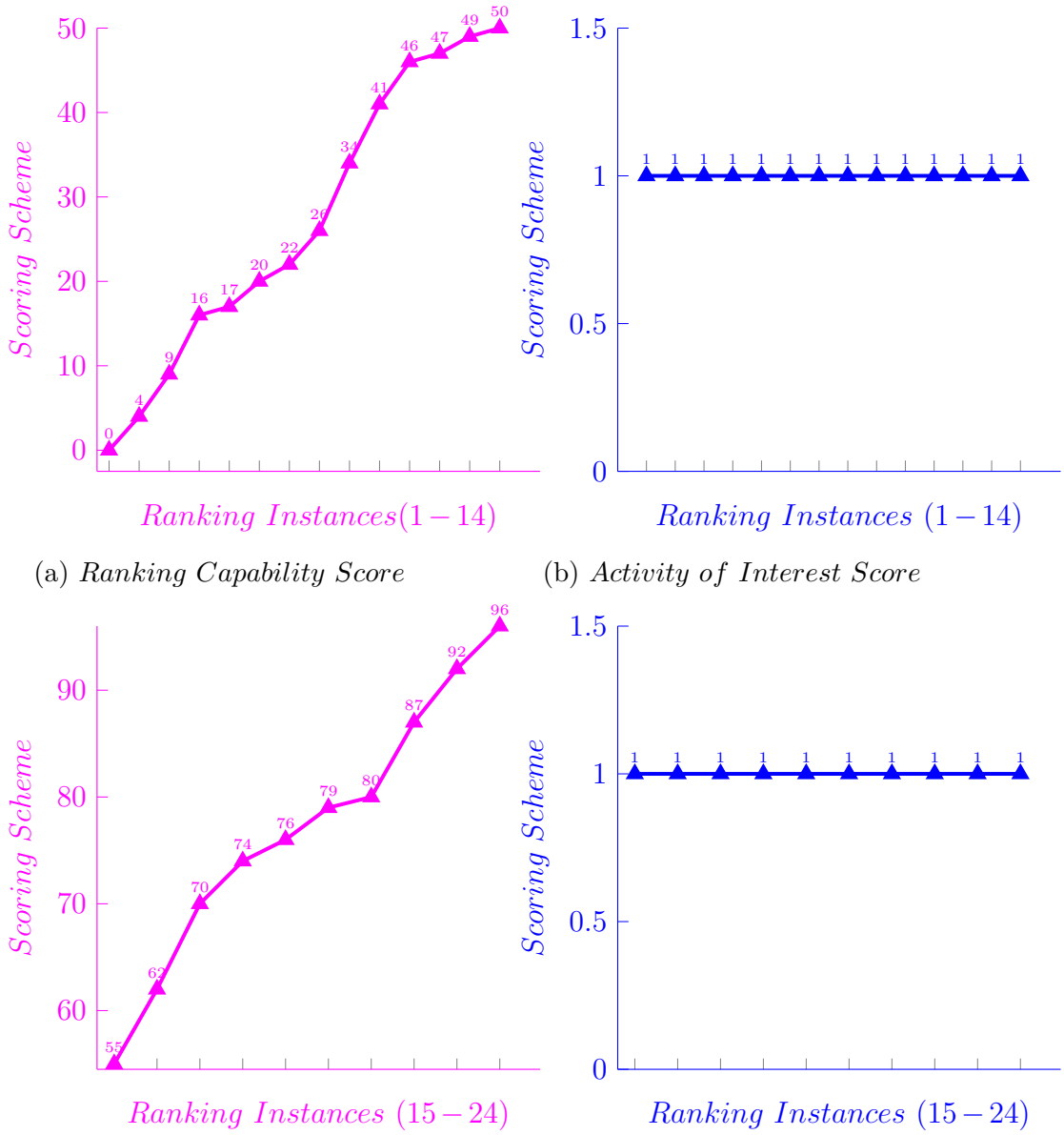
On the other hand, the RCS provides unique scores for all the ranking instances in three different scenarios as it is shown in Figures 3.5, 3.6, 3.7 parts (a) (c). Similarly, in the first scenario, the first ranking instances score 0 indicating that the SWA system has ranked the emerging situations as perfect as the ground truth. While the second ranking instances scores (1) indicating that the SWA system has ranked the emerging threats as opposed to the ground truth. While in the second and third scenario in Figures 3.6, 3.7 the (RCS) provide unique scores for all the ranking instances over the underlying situational assessment. This is because it has been designed to evaluate ranking capability in relation to the prioritisation process, specifically under situational assessment where the SWA system is reporting only distinct threaten situations but with different severity paradigms.

The proposed performance metric can provide another dimensional support for the decision making resources. Likewise, it can direct the analyst's attention about any ranking capability issues which may occur during real time operation.

3.6 Conclusion

This chapter has examined the performance metric, namely the Activity of Interest Score $AoIScore$, under an extended scenario where the SWA system has identified only important activities but with different severity paradigms. The results showed that the $AoIScore$ did not quantify the ranking capabilities for different levels of assessment appropriately. Rather, it treated all the identified prioritised events only as important activities, regardless of their respective degree of importance. In regards to this, section 3.3.3 introduced a new performance metric for evaluating the ranking capability for the underlying scenario.

To evaluate the underlying performance metrics against their intended purpose, we examined each scoring scheme with the help of a case study. The obtained results



(a) *Ranking Capability Score* (b) *Activity of Interest Score*

(c) *Ranking Capability Score* (d) *Activity of Interest Score*

Fig. 3.7 Quality based evaluation for RCS versus $AoIScore$. $(4!) = 24$ *Ranking Instances*

showed an evidencing score for only two or three ranking instances. However, to examine the proposed performance metrics against all the ranking instances, we conducted a quality-based evaluation. The evaluation process examined the scalability

of the proposed scoring schemes against all the ranking instances over three separates scenarios; each encompassed a different number of prioritised events.

The outcome of the evaluation process was as follows: the *AoIScore* didn't provide an appropriate scoring scheme against the extended situational assessment. Furthermore, it scored (1) against all the ranking instances for three different scenarios. This is because it was not designed to evaluate the ranking capability in relation to the prioritisation process; rather it was intended to evaluate the scheduling process of a real time system where the underlying system is reporting two different classes of tracking activity. The first and second class are activities with significant impact and activities with minimal impact on the dynamically monitored environment, respectively.

Alternatively, the Ranking Capability Score has provided an appropriate scoring scheme against all the ranking instances over three separates scenarios. Indeed, from the data fusion perspective, it successfully quantified the ranking capabilities in relation to the prioritisation process. In particular, under situational assessment where the SWA system identified (or configured to report) only important activities but with different severity paradigms.

The proposed performance metric can provide another dimensional support for the decision making resources. Specifically it provides an indications scoring scheme about three qualitative states that are likely to occur in a real-time operation. These are *Good State*, *Degraded State* and *Bad state* where the SWA system is ranking the identified activities as perfect as the ground truth, as not as perfect as the ground truth, and as opposed to the ground truth, respectively. Furthermore, it can direct the analyst's attention about any ranking capability issues that may occur during a real time operation.

Chapter 4

Enhanced Ranking Capability Score′(RCS')

4.1 Introduction

The data fusion community [27], [10],[26], [11]has introduced multiple procedures of situational assessments to facilitate timely responses for emerging situations. Process Refinement (level 4) of the Joint Directors of Laboratories (JDL) [53], [54], [82] is a meta-process used to assess and improve the data fusion task during real time operations. The previous chapter introduced the *Ranking Capability Score′* (RCS') for evaluating the prioritisation process of a real time system. However, the process of user refinement (level 5) of the JDL is intended to address knowledge representation [11], [7], [8],[11], [7], [8], for the decision-making resources.

We have proposed a performance metric [79] to evaluate the prioritisation process of situational awareness domains. Furthermore, we examine the proposed performance metrics with two levels of assessment. The first level is a case study based evaluation[79] which is used to guide researchers from various disciplines on adapting the Ranking Capability Score (RCS) into their domain specific solution. The second level is a quality

based evaluation[80] used to examine the proposed performance metrics against their intended purpose. In both cases, the RCS has successfully quantified the prioritisation process of a real time system.

This chapter introduces a third level of assessment in order to examine the reliability of the proposed scoring scheme against the cognitive states of user perception. The evaluation process encompasses two phases. The first phase examines the RCS against the three qualitative states that are likely to occur in a real time operation. These are the *Good State*, *Degraded State* and *Bad State* where the SWA system ranks the identified activities as perfect as the ground truth, respectively. The second phase examines the scoring schemes further with the help of a case study based scenario to illustrate the knowledge representation problem between the RCS and user perception.

We need to know what is the usefulness of the underlying metrics in directing the analyst team towards any ranking capability issues during a real time operation, concerning different domains such as air, sea, land, space and cyberspace. To answer the underlying research question, this project has conducted the following objectives. We reviewed existing research efforts [6][39] [65] [75] which provided ranking methods based on various criteria. We selected a case study [75] where the author discussed a number of a logarithm and techniques over the past 10 years, for enabling the situational awareness domain (perception, comprehension and projection) as well as filtering and ranking different tracking activities during a real time operation. Next, we introduced a ranking capability problem for the selected scenario; this was to allow the RCS [79] to evaluate different ranking instances concerning the three qualitative states mentioned above. Finally, we observed the obtained scoring scheme versus the user perception to illustrate the knowledge representation problem during a real time operation.

This chapter is divided as follows. The first section examines the " Ranking Capability Score" (RCS) versus the three qualitative states mentioned previously; *Good State*, *Degraded State* and *Bad State*. The second section introduces an enhanced method for the proposed metrics, called the enhanced *Ranking Capability Score′* (RCS'). The third section conducts a reliability based evaluation with the help of a case study based scenario, to evaluate the proposed enhanced scoring scheme against the operator perception. The fourth section presents a comparative evaluation and, finally, we discuss our findings and future work.

4.2 Knowledge Representation Problem for the Ranking Capability Score

This section examines the " Ranking Capability Score" (RCS) versus the three following qualitative states, *Good State*, *Degraded State* and *Bad state* where the SWA system is ranking the identified activities as perfect as the ground truth, not perfect as the ground truth, and opposed to the ground truth, respectively.

This section is divided as follows. The first part illustrates the knowledge representation problem concerned with the RCS and the second part further explains the knowledge representation problem with the help of a case study from the Cyberspace domain. The third part presents an extended scenario from the demonstrated case study focusing on when the system is experiencing ranking capability issues during the multi level SWA. Furthermore, the reliability issues of the RCS are discussed in terms of providing unreliable scores for the three qualitative states previously described, specifically, when the proposed number of tracking activities are dynamically changing in the multi level situational assessment.

4.2.1 Number of Tracking Activity versus Qualitative States

In the light of the three qualitative states previously mentioned, this section investigates the reliability of the ranking capability scoring scheme over three separate scenarios. Each scenario encompasses a different number of prioritised events. The reliability based evaluation investigates the knowledge representation issue of the RCS. It is intended to answer the following research question; to what extent can the underlying scoring scheme be helpful in directing the analyst's attention to any ranking capability issues during a real time operation, specifically, where the proposed number of tracking activities are dynamically shifted in a multi projection environment.

The evaluation process for the first scenario observes the ranking capability scoring scheme when the real time system is reporting at least two events.

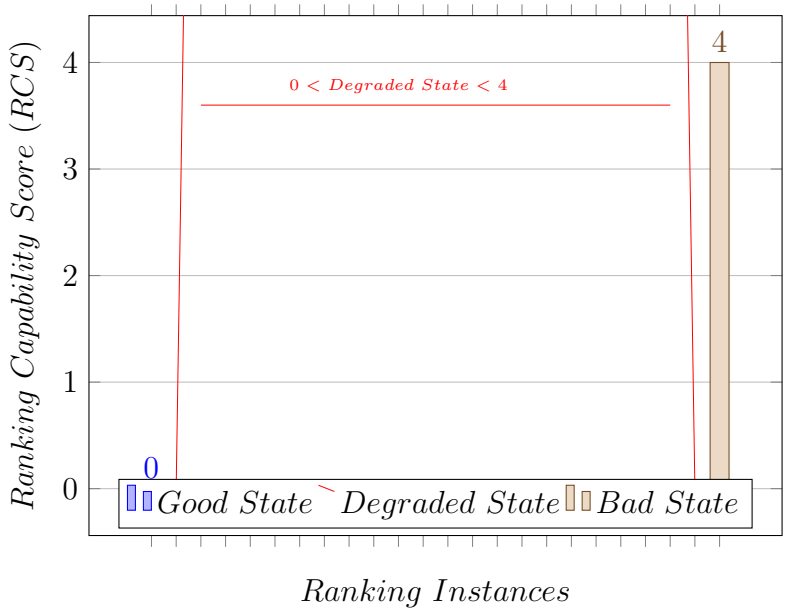


Fig. 4.1 Reliability Based Evaluation for the *Ranking Capability Score* (RCS) (2!) 2 Ranking Instances

The obtained result for the first scenario is shown in Figure 4.1; the proposed number of emerging events are equal to two tracking activities of interest (AoIs), ($N = (2)$). The scoring scheme for the three qualitative states is shown in Figure 4.1.

According to the proposed number of tracking activities (2!), there are only two ranking instances. The first and second ranking instances represent the *Good State* and *Bad State* where *RCS* scores are (0) and (4), respectively.

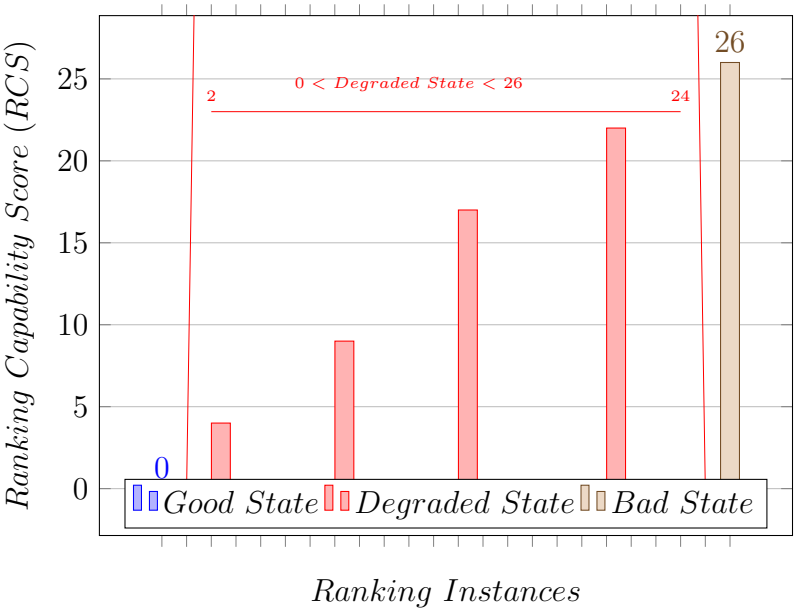


Fig. 4.2 Reliability Based Evaluation for the *Ranking Capability Score* (*RCS*) (3!) 6 Ranking Instances

The second scenario observed the RCS when the real time system was reporting at least three tracking activities; results are shown in Figure 4.2. The number of emerging events is equal to three tracking activities of interest (*AoI*s), ($N = (3)$). The scoring scheme for the three qualitative states is shown in Figure 4.2. According to the number of prioritised events (3!), there are only (6) ranking instances for the underlying scenario. The first ranking instances represent the *Good State*, where *RCS* score (0). The last ranking instance represents the *Bad State* and it scores (26). The ranking instances from ((2)*to*(5)) represent the *Degraded State*. The *RCS* scores ((2)*to*(24)) over all the ranking instances for the underlying state respectively.

The third scenario observed the RCS when the real time system was reporting at least four tracking activities; results are shown in Figure 4.3. The number of emerging

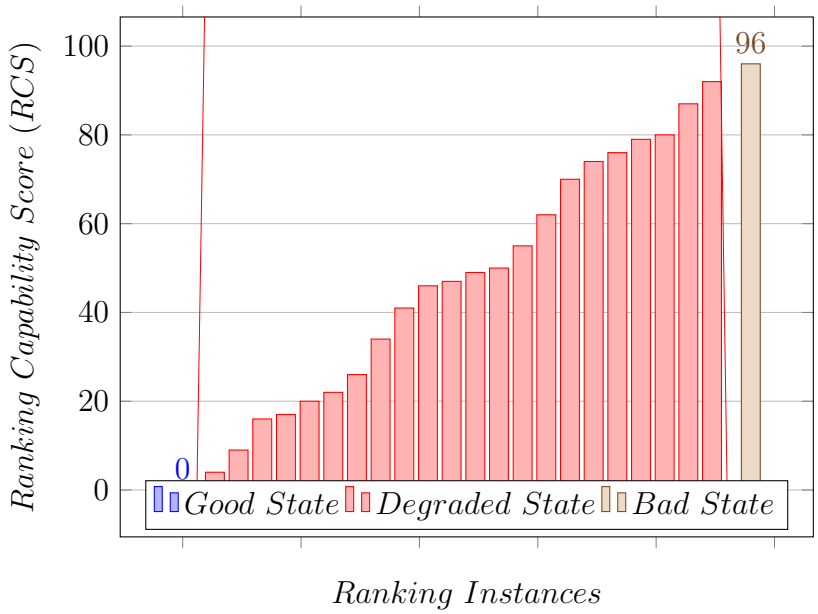


Fig. 4.3 Reliability Based Evaluation for the *Ranking Capability Score* (RCS) (4!) 24 Ranking Instances

events are equal to four tracking activities of interest (AoIs), ($N = (4)$). Similarly, the scoring scheme for the three qualitative states are shown in Figure 4.3. According to the number of prioritised events (4!), there are only (24) ranking instances for the underlying scenario. the first and last ranking instances represent the *Good State* and *Bad State*, where the RCS score was (0) and (96), respectively. The ranking instances from ((2)*to*(23)) represent the *Degraded State* where the RCS scored (2) to (92) for all the ranking instances of the underlying qualitative state.

Regrettably, the underlying performance metric, RCS, for the three qualitative state changes in relation to the proposed number of emerging activities, and did not provide reliable scores for the three qualitative states.

The *Bad State* scored 4 for the first scenario (Figure 4.1), 26 for the second scenario (Figure 4.2) and 96 for the last scenario (Figure 4.3). Moreover, the dimensions of the degraded states also changed over the three scenarios. When the proposed number of prioritised events are either increased or decreased, at any given instance the RCS

for the three qualitative states also changed accordingly. Hence, the scoring scheme inherited a knowledge representation problem, and it can easily degrade the cognitive status of the user perception during a real time operation.

The next section further examines the underlying scoring scheme, RCS, with the help of a case study, to guide the researchers from various areas, how to adopt the proposed metrics to their domain specific needs, and to evaluate the proposed scoring scheme against the operator perception, The evaluation process intends to validate the scoring scheme against the decision-making perception.

4.2.2 Case Study: Multi Projection Environment

This section conducts a reliability based evaluation for the RCS, with the help of a case study based scenario. We need to know to what extent the underlying metric can be helpful in directing the analyst team to any ranking capability issues, during a real time operation in different domains such as air, sea, land, space and cyberspace.

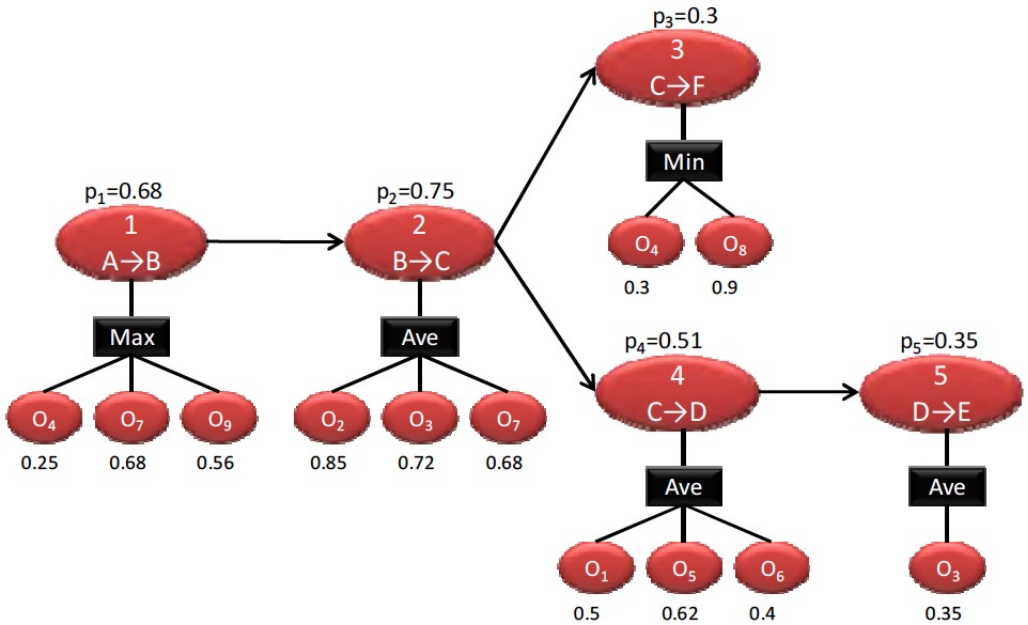


Fig. 4.4 Ranking Paradigams For The Identified Events Based on Two Levels of Assessment(adapted from [75])

In the light of existing literature, a case study has been demonstrated [75] [74] where authors applied multiple perspective measures against the emerging situation. It is important to note that the authors emphasised mathematical techniques rather than representing whole system implementation. The first level of assessment identified five intrusion activities, shown in Figure 4.4.

The second level of assessment applied a contextual perspective measure against the identified situation. This is to rank the most important activity into a predetermined order. As a result, the ranking paradigms for the comprehension stage differed to the perception stage. The figure (4.4) presents two ranking paradigms against the identified tracking activity.

The first ranking paradigm relates to the first level of assessment for the perception stage. The system identified five intrusions activities. The first ($A \longrightarrow B$) scored (1) while the second one ($B \longrightarrow C$) scored (2). The third one ($C \longrightarrow F$) scored (3), the fourth one ($C \longrightarrow D$) scored (4) and finally, the fifth one ($D \longrightarrow E$) scored 5).

The second ranking paradigm relates to the second level of assessment for the comprehension stage where the underlying process ranked the identified activities based on their current impact. The first intrusion activity ($A \longrightarrow B$) scored (0.68). While the second one ($B \longrightarrow C$) scored (0.75). The third one ($C \longrightarrow F$) scored (0.3). The fourth one ($C \longrightarrow D$) scored (0.51) and finally the fifth one ($D \longrightarrow E$) scored (0.35).

Noticeably, the automation process of information perception, projection and comprehension encompassed two simultaneous operations: filtering and prioritisation. Hence, the number of tracking activities proposed by the situational awareness domain can dynamically change. Furthermore, the higher level of data fusion (projection stage) focusses on the threats which affect the outcome of things that "will" happen. Therefore, the system at the third level of assessments has anticipated four intrusion activities. In Figure 5.6 the underlying process has used different pieces of information

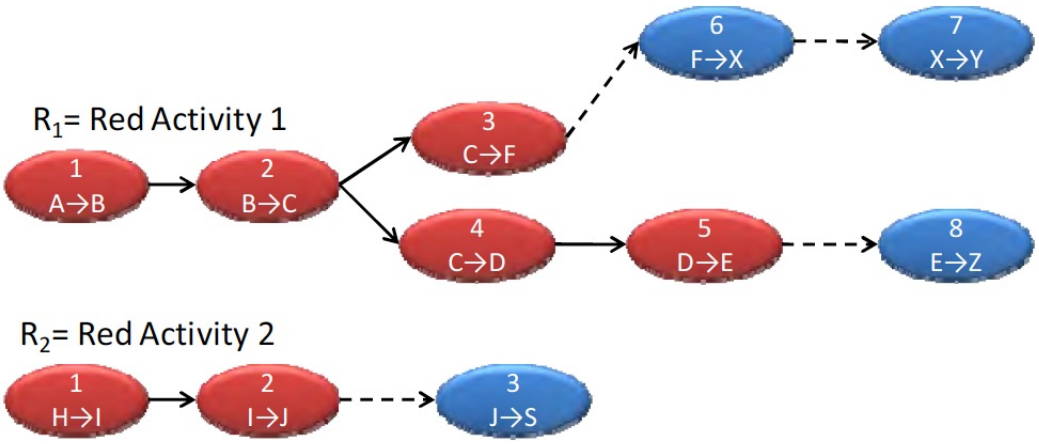


Fig. 4.5 Filtering Out The Identified Events Based On Their Current Impacts and Future Threats (adapted from [75])

about the protected domain and the emerging intrusion activity to estimate future threats for each tracking activity. The ranking paradigms for these events are shown in Figure 4.6

	Plausible Threatened Assets				**Weight**
	6: F→X	**7: X→Y**	**8: E→Z**	**3: J→S**	
f_1	0.1	0.6	0.6	0.9	1
f_2	0.1	0.2	0.5	0.9	2
f_3	0.1	0.8	0.6	0.01	4
f_4	0.1	0.4	0.4	0.9	1
T_a	**0.50**	**0.85**	**0.87**	**0.49**	

Fig. 4.6 Ranking Paradigms for the Third Level of Assessments (adapted from [75]

The first intrusion activity ($F \longrightarrow X$) scored (0.50) while the second ($X \longrightarrow Y$) scored (0.85). The third ($E \longrightarrow z$) scored (0.87) and finally, the fourth ($J \longrightarrow S$) scored (0.51).

Based on the demonstrated case study, we can conclude that the number of emerging activities are dynamically changing during a real time operation. In light of the above situational assessments, the perception (first level of assessment) and comprehension

stages (second level of assessment) have a different number of threats in comparison to the projection stage.

In regards to this, in the next section we propose an extended scenario from the demonstrated case study to examine the RCS against the user perception.

4.2.3 Extended Scenario: Inline Situational Assessment

This section does not change the original situational assessment for the multi-projection environment. Rather, it considers a corner case from the demonstrated case study.

We are interested in evaluating the ranking capability for a multi-projection environment, specifically during a real time operation; it is expected at any point in time to verify the ranking capability for each level of assessments.

In light of the case study demonstrated previously, the evaluation process will quantify the prioritisation process for the first level of assessment. In particular, when the perception stage has identified only two intrusion activities. The ground truth for the perception stage is shown in Table 4.1.

Table 4.1 Pre-determined Order for Identified Threat at the Perception Stage

Ground Truth	**Activity of Interest**	**Priority**
GT_0	(B$\rightarrow C$)	1
GT_1	(A$\rightarrow B$)	2

The underpinning assumption for the proposed ranking paradigms, shown in Table 4.1, is based on the intrusion progress inside the protected environments. Likewise, the intrusion events $(B \rightarrow C)$ are advancing inside the protected network deeper than $(A \rightarrow B)$.

Furthermore, the intrusion activity $(B \rightarrow C)$ has a higher credibility score in comparison to $(A \rightarrow B)$ and therefore, the $(B \rightarrow C)$ has the highest likelihood score

with the priority value of (1), whereas stepping $(A \rightarrow B)$ has the least credibility with a priority value of (2).

Additionally, the evaluation process is expected to evaluate the ranking capability for the second level of assessment, specifically where the comprehension stage has identified three intrusion activities. The absolute truth ranking paradigms are shown in Table (4.2).

Table 4.2 Pre-determined Order for Identified Threat at the Comprehension Stage

Ground Truth	**Activity of Interest**	**Priority**
GT_0	(B$\rightarrow C$)	1
GT_1	(A$\rightarrow B$)	2
GT_2	(C$\rightarrow F$)	3

The credibility value for the proposed threats in Table 4.2 is inspired by the ranking paradigms based on the case study demonstrated in section 4.2.2; this is shown in Figure (4.4). The intrusion activities $(B \rightarrow C)$ have the highest credibility score with a priority value of (1) while the attacker steps $(A \rightarrow B)$ have a lesser credibility score with the priority value of (2). Moreover, the third intrusion activities $(C \rightarrow F)$ had the least credibility score with a priority value of (3).

Finally, the evaluation process evaluates the ranking capability for the projection stage specifically where the SWA system has anticipated four threatening activities. The ground truth for the projection stage is shown in Table 4.3.

Table 4.3 Pre-determined Order for Identified Threat at the Projection Stage

Ground Truth	**Activity of Interest**	**Priority**
GT_0	(E$\rightarrow Z$)	1
GT_1	(X$\rightarrow Y$)	2
GT_2	(F$\rightarrow X$)	3
GT_3	(J$\rightarrow S$)	4

The credibility value for the anticipated threats (shown in Table 4.3) are inspired by the ranking paradigms based on the case study demonstrated in section 4.2.2. This is shown in Figure (4.6).

The third level of assessment applied a contextual perspective measure against the emerging situation. In return, the projection stage anticipated four intrusion activities (shown in Figure 4.6); each one has different severity paradigms (shown in Table 4.3.

The $(E \rightarrow Z)$ had the highest credibility score with the priority value of 1 while the second and third intrusion activities, $(X \rightarrow Y)$ and $(F \rightarrow X)$ had fewer credibility scores with priority values of 2 and 3, respectively. However, the $(J \rightarrow S)$ had the least credibility score with a priority value of 4.

This section has introduced the absolute ranking paradigms for the extended scenario. The next stage presents the proposed ranking instances for each level respectively. In other words, the evaluation process will introduce a ranking capability issue for each projection technique randomly.

The random assumption does not impact the underlying evaluation since the assessment process is not interested in the capability of a particular system "yet". Rather, it intended to examine the proposed performance metric against the cognitive status of user perception, in particular during a real time operation, where the number of emerging activities shifts in multi-level situational assessments.

Hence, the random assumption of the ranking paradigms for each projection technique can serve the reliability based evaluation. In regard to this, we have assumed three different levels of assessments. The proposed assessment is for the perception stage (provided in Table 4.4). The ranking paradigms for the first projection stage is inherited from the following assumption [77] where the perception stage is reporting the emerging threats based on their time of occurrence.

Table 4.4 Proposed Assessment at the Perception Stage (Level 1)

Proposed Assessment	**Activity of Interest**	**Priority**
PA_0	(B$\rightarrow C$)	1
PA_1	(A$\rightarrow B$)	2

Moreover, the proposed assessment for the comprehension stage (shown in Table 4.4) has ranked the activity as opposed to the ground truth.

Table 4.5 Proposed Assessment at the Comprehension Stage (Level 2)

Proposed Assessment	**Activity of Interest**	**Priority**
PA_0	(C$\rightarrow F$)	3
PA_1	(A$\rightarrow B$)	2
PA_2	(B$\rightarrow C$)	1

Furthermore, the proposed assessment for the projection stage (shown in Table 4.6) has ranked the anticipated threats as not perfect as the ground truth.

Table 4.6 Proposed Assessment at the Projection Stage (Level 3)

Proposed Assessment	**Activity of Interest**	**Priority**
PA_0	(E$\rightarrow Z$)	1
PA_1	(J$\rightarrow S$)	4
PA_2	(F$\rightarrow X$)	3
PA_3	(X$\rightarrow Y$)	2

The demonstrated case study helps us to examine the reliability of *RCS* scoring scheme against the three qualitative states mentioned previously; the *Good State*, *Bad State*, and *Degraded State*. The next section quantifies the ranking capability for the underlying scenario, using the proposed performance metric, the Ranking Capability Score (RCS).

4.2.4 Measuring the Ranking Capability of a Real Time System

As previously explained, any situational assessment encompasses two components: the proposed assessment outputs and its ground truth paradigms. We have defined analytically six different objects, concerning the prioritisation process of a real time system; three objects for the proposed assessment components and another three objects for the ground truth components.

According to the in-line situational assessment, demonstrated in section 4.2.3, there are three *proposed assessment* outputs concerning three different ranking instances for the identified situation. Each level of assessment has a reference ranking paradigm for the identified events. This is called the *ground truth*.

To assess the ranking capability for each assessment level, the evaluation process will first substitute relevant values for the absolute truth ranking paradigms and the *proposed assessment* outputs, concerning the multilevel situational assessments for the demonstrated case study in section 4.2.2. The second phase is to quantify the ranking paradigms for each level using the RCS, the quantitative assessment aimed at verifying the proposed assessment ranking instances in comparison to their absolute truth ranking paradigms.

This section is divided as follows. The first section illustrates the substitution process for extracting relevant values for the ground truth objects and the second section demonstrates the substitution process concerning the proposed assessments outputs. Finally, the third section provides a guide for researchers on how to substitute the situational assessment components in order to quantify the ranking capability of a real time system.

4.2.5 Ground Truth Paradigms

This section illustrates the substitution of three objects concerning the ground truth component; they have been defined in equation 3.6 Therefore, the first object is the AoI_i (defined in equation 3.3), while the second object is SoI_j (defined in equation 3.5), and the third object is NoH_i (defined in equation 3.4).

This section is divided as follows: the first section illustrates the substitution process concerning the reference ranking paradigms for the perception stage outputs (shown in Table 4.1). The second section demonstrates the substitution process for the comprehension stage (shown in Table 4.5), and, finally, the third section explains the projection stage (shown in Table 4.3).

Perception Stage Level 1

This section substitutes relevant values from the ground truth output (shown in Table 4.1). The first object is the activity of interest $AoI_i = \{aoi_1(1), aoi_2(2)\}$, where $N = 2$. The second object is the score of importance $SoI_j = \{soi_1(2), soi_2(1)\}$ and the third object is the predetermined order of each complex event $NoH_i = \{noh_1(1), noh_2(0)\}$, where $ToH = 1$.

Comprehension Stage Level 2

This section substitutes relevant values from the ground truth output (shown in Table 4.5). The first object is the activity of interest $AoI_i = \{aoi_1(1), aoi_2(2), aoi_3(3)\}$, where $N = 3$. The second object is the score of importance $SoI_j = \{soi_1(3), soi_2(2), soi_3(1)\}$ and the third object is the predetermined order of each complex event $NoH_i = \{noh_1(2), noh_2(1), noh_3(0)\}$, where $ToH = 2$.

Projection Stage Level 3

This section substitutes relevant values from the ground truth output (shown in Table 4.3). The first object is the activity of interest $AoI_i = \{aoi_1(1), aoi_2(2), aoi_3(3), aoi_4(4)\}$, where $N = 4$. The second object is the score of importance $SoI_j = \{soi_1(4), soi_2(3), soi_3(2), soi_4(1)$ and the third object is the predetermined order of each complex event $NoH_i = \{noh_1(3), noh_2(2), noh_3(1), noh_4(0)\}$, where $ToH = 3$.

4.2.6 Proposed Assessment Outputs

This section illustrates the substitution of three objects concerning the proposed assessment outputs; they have been defined in (equation number 3.9) first object is the proposed ranking for the activity of interest AoI_r (defined in equation 3.7) The second object is the score of importance SoI_r (defined in equation 3.8) and the third object is the predetermined order based on the number of hops NoH_i (defined in equation 3.4).

This section is divided as follows: the first section illustrates the substitution process concerning the proposed ranking paradigms for the perception stage outputs (shown in Table 4.4). The second section demonstrates the substitution process for the comprehension stage (shown in Table 4.5) and, finally, the third section explains the projection stage (shown in Table 4.6).

Perception Stage Level 1

This section substitutes relevant values from the proposed assessment outputs of the perception stage (shown in Table 4.4). The first object is the activity of interest $AoI_r = \{aoi_r(1), aoi_r(2)\}$, where $N = 2$. The second object is the score of importance $SoI_r = \{soi_r(2), soi_r(1)\}$. The third object is the predetermined order of each complex event $NoH_i = \{noh_1(1), noh_2(0)\}$, where $ToH = 1$.

Comprehension Stage Level 2

This section substitutes relevant values from the proposed assessment output of the comprehension stage (shown in Table 4.5). The first object is $AoI_r = \{aoi_r(3), aoi_r(2), aoi_r(1)\}$. The second object is the score of importance $SoI_r = \{soi_r(1), soi_r(2), soi_r(3)\}$. The third object is the number of hops, $NoH_h = \{noh_1(2), noh_2(1), noh_3(0)\}$.

Projection Stage Level 3

This section substitutes relevant values from the proposed assessment output of the projection stage (shown in Table 4.6).
The first object is $AoI_r = \{aoi_r(1), aoi_r(4), aoi_r(3), aoi_r(2)\}$.
The second object is the score of importance $SoI_r = \{soi_r(4), soi_r(1), soi_r(2), soi_r(3)\}$.
The third object is the number of hops, $NoH_h = \{noh_1(3), noh_2(2), noh_3(1), noh_4(0)\}$.
This section has substituted the situational assessment components for the multilevel situational assessments. The next section illustrates the second phase for the evaluation process.

4.2.7 Quantification Process

This section illustrates the evaluation process for quantifying the ranking capability of a real time system. This section is divided as follows: the first section quantifies the proposed assessment ranking paradigms of the perception stage (shown in Table 4.4). The second section quantifies the proposed assessment ranking instances for the comprehension stage [shown in Table 4.5]. Finally, the third section illustrates the method of quantifying the proposed ranking instances for the projection stage(shown in Table 4.6).

Perception stage

The evaluation process quantifies the ranking capability for the first level of assessment. The first part quantifies the first projection techniques, where the perception stage has identified two AoI. This is done by evaluating the proposed assessment ranking instances (shown in Table 4.4) in comparison with the absolute truth ranking paradigms of the situation (provided in Table 4.1).

Firstly, the evaluation process computes the current number of hops (defined in equation 3.10) $CoH_i = \{coh_1(2-1=1), coh_2(2-2=0)\}$. Secondly, it computes the actual number of hops, (defined in equation 3.11) $AoH_i = \{aoh_1(1-0=1), aoh_2(1-1=0)\}$. Thirdly, it computes the difference in the number of hops, (defined in equation 3.12, $DoH_j = \{doh_1(0), doh_2(0)\}$. Fourthly, it computes the reference number of hops for each AoI (defined in equation 3.13 $RoH_i = \{roh_1(2*0=0), roh_2(1*0=0)\}$. Finally, it quantifies the ranking capabilities using the *Ranking Capability Score* for the perception stage as follows: $RCS = \sum_{j=1}^{2} = roh_1\,(0) + roh_2\,(0+0) = 0$.

Comprehension Stage

The evaluation process quantifies the ranking capability for the second projection technique. The second section quantifies the second projection technique, where the comprehension stage has proposed three AoI. This is done by evaluating the proposed assessment (shown in Table 4.5) in comparison to the absolute truth of the situation (provided in Table 4.2).

Firstly, it computes the current number of hops, (defined in equation 3.10 $CoH_i = \{coh_1(3-3=0), coh_2(3-2=1).coh_3(3-1=2)\}$. Secondly, it computes the actual number of hops (defined in equation 3.11 $AoH_i = aoh_1(2-0=2)$, $aoh_2(2-1=1).aoh_3(2-2=0)$. Thirdly, it computes the difference in number of hops (defined in equation 3.12, $DoH_j = \{doh_1(2), doh_2(2).doh_3(0)\}$. Fourthly, it computes

the reference number of hops for each AoI,(defined in equation 3.13, $RoH_i = \{roh_1(3*2=6), roh_2(2*2=4), roh_3(1*0=0)\}$. Finally, it quantifies the ranking capabilities using the *Ranking Capability Score* for the comprehension stage as follows: $RCS = \sum_{j=1}^{3} = roh_1(6) + roh_2(6+4) + roh_3(6+4+0) = 26$.

Projection Stage

The evaluation process quantifies the ranking capability for the third projection technique. In the third section the RCS aimed to quantify the third projection technique, where the projection stage has anticipated four AoI. This is done by verifying the proposed assessment ranking instances (shown in Table 4.6) in comparison to its absolute truth ranking paradigms (shown in Table 4.3).

Firstly, the underlying process computes the current number of hops,(defined in equation 3.10 $CoH_i = \{coh_1(4-1=3), coh_2(4-4=0), coh_3(4-3=1),$
$coh_4(4-2=2)$. Secondly, it computes the actual number of hops (defined in equation 3.11) $AoH_i = \{aoh_1(3-0=3), aoh_2(3-1=2).aoh_3(3-2=1), aoh_3(3-3=0)\}$. Thirdly, it computes the difference in number of hops (defined in equation 3.12), $DoH_j =$
$\{doh_1(0), doh_2(2).doh_3(2), doh_4(0)\}$. Fourthly, it computes the reference number of hops for each AoI, (defined in equation 3.13, $RoH_i =$
$\{roh_1(4*0=0), roh_2(3*2=6), roh_3(2*2=4), roh_4(1*0=0)\}$. Finally, it quantifies the ranking capabilities for the projection stage, using the *Ranking Capability Score* as follows: $RCS = \sum_{j=1}^{4} = roh_1(0) + roh_2(0+6) + roh_3(0+6+4) +$
$roh_4(0+6+4+0) = 26$.

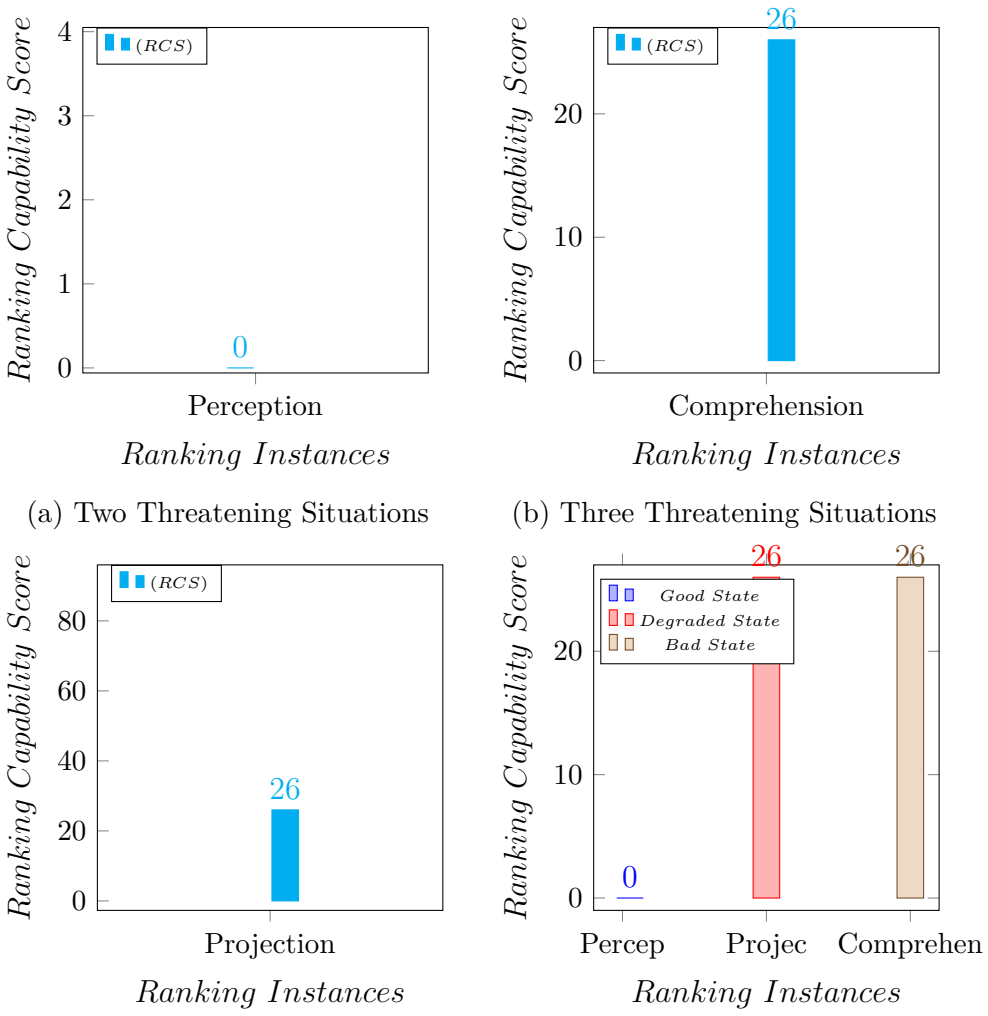
4.2.8 Assessment Results

This section presents the obtained results from the reliability based evaluation. The discussion will encompass two levels of assessments. At first glance, Figure 4.7-(a),(b) and (c) presented the RCS against each level of assessment, in relation to the operator cognitive status. However, at second glance, Figure 4.7 part (d) demonstrates the returned scoring scheme, versus the three qualitative states.

The first level of evaluation has three points of comparison. The first is presented in Figure 4.7-(a). The first level of assessment scores (0). This means, the perception stage has ranked two intrusion activities (shown in Table 4.4) as perfect as the ground truth (shown in Table 4.1). Hence, the dimension of the RCS ,in relation, to the number of identified threats, are shown between (0) to (4). From these, (4) presented the worst state for the underlying situational assessment.

The second point of comparison is shown in Figure 4.7-(b). The second level of assessments has scored (26). This means, the comprehension stage has ranked three intrusion activities (shown in Table 4.5) as opposed to the ground truth (shown in Table 4.2). Hence, the dimension of the RCS, in relation to the number of prioritised events, are shown between (0) and (26). From these, (0) presented the good state for the underlying situational assessment while (26) represented the bad states.

The third point of comparison is shown in Figure 4.7-(c). The third level of assessment scores (26). This means, the projection stage has ranked four anticipated threats (shown in Table 4.6) as not perfect as the ground truth (shown in Table 4.3). Hence, the dimension of the RCS , in relation to the number of anticipated threats, is shown between (0) and (96). From these, (0) presented the good state, while (96) represented the worst state for the underlying situational assessment.



(a) Two Threatening Situations (b) Three Threatening Situations

(c) Four Threatening Situations (d) Multi Projection Environment

Fig. 4.7 Measuring the Ranking Capability for Multi Projection Environment Using the *Ranking Capability Score* (*RCS*)

The RCS has quantified three different ranking instances appropriately. However, from the cognitive status of user perception, the underlying performance metric has provided unreliable scores for two distinct qualitative states.

Figure 4.7-(*d*) presented the RCS versus the three qualitative states. The *Good State* scores (0) for the first projection technique, in which the perception stage has ranked two intrusion activities as perfect as the ground truth. The *Degraded State* scores (26) for the second level of assessments, in which the projection stage has ranked

four anticipated threats as not perfect as the ground truth. While the *BadState* also scores (26) for the third level of assessments, where the comprehension stage has ranked three intrusion activities as opposed to the ground truth.

Apparently, the ranking paradigms for the two qualitative states are different(degraded states and bad states). However, the (RCS) provides a similar scoring scheme for both stages. This is because the underlying performance metric was originally designed for a different situational assessment in which the number of prioritised events are fixed. In other words, the scoring scheme evaluates the prioritisation process under a contextual scenario, where the real time system is configured to perform only a prioritisation process, against the emerging situation. Specifically, the number of emerging events for the perception stage is the same as the comprehension and projection stages.

The next section presents the enhanced Ranking Capability Score′ (RCS') to provide reliable scoring scheme for the underlying situational assessment, where the real time system is configured to perform filtering and prioritisation processes during real time operation. Consequently the proposed number of tracking activity are dynamically shifted in multi projection environment.

4.3 Enhanced Method of the Ranking Capability Score

This section introduces the Ranking Capability Score (RCS'). This section is divided as follows: the first section demonstrates the development phase for the underlying performance metric. The second section presents the reliability based evaluation in order to quantify the ranking capability for the demonstrated case study in section 4.2.3. Finally, the third section illustrates the obtained results.

4.3.1 Ranking Capability Score′(RCS′)

This section discusses the development phase for *Ranking Capability Score′* (RCS'). This underlying metric enhances the scoring scheme of the "RCS" , in term of providing a representative scores for the operator perception concerning the three qualitative states mention above, and therefore should provide an appropriate scoring scheme for the prioritisation process during real time operation, in particular, under a scenario where the number of tracking activity are dynamically changed in real time operation.

In order to enhance the RCS, the underlying process aimed to normalise the scoring scheme between (0) and (1). With this in mind, the development phase is divided into two stages. The first stage introduces three elements that are preliminary for the the normalization process.

The first element, is the reverse set of AoH_i and it is defined in 4.1:

$$AoH'_j = ToH - NoH_i \tag{4.1}$$

where

The total number of hop $ToH = N - 1$, $1 \leq i \leq N$, $1 \leq j \leq N$

The second object is the worst state for the DoH_j, it is defined in 4.2:

$$DoH'_j = \sum_{i=1}^{j}(NoH_i) - \sum_{i=1}^{j}(AoH'_i) \tag{4.2}$$

where

$1 \leq i \leq N$, $1 \leq j \leq N$

The third object introduces the worst state for the RoH_i and it is defined in 4.3:

$$RoH'_i = SoI_i * DoH'_i \tag{4.3}$$

where

$1 \leq i \leq N$ And $1 \leq j \leq N$

The process is satisfactory after introducing all the complementary objects for the first development process. Next, the underlying process demonstrates the normalization phase for the RCS; it is defined in equation 4.4). We call it the *Ranking Capability Score′*.

$$RCS' = \frac{\sum_{j=1}^{N} \left(\sum_{i=1}^{j} (SoI_i)\, DoH_i \right)}{\sum_{j=1}^{N} \left(\sum_{i=1}^{j} (SoI_i)\, DoH'_i \right)} \tag{4.4}$$

where, $1 \leq i \leq N \quad 1 \leq j \leq N$

This section introduced the development phase for the underlying performance metric RCS'. The next section examines the proposed performance metrics against their intended purpose, with the help of the case study demonstrated in 4.2.2.

4.3.2 Measuring the Ranking Capability of a Real Time System

According to the in-line situational assessment (demonstrated in 4.2.3) there are three different *ground truths* for the multi projection environment. Each one represents the absolute truth ranking paradigms, in relation to the perception, comprehension and projection stages, respectively.

Moreover, there are three different levels of assessment against the emerging situation. The first level is the *proposed assessment* for the perception stage. The second level is the *proposed assessment* for the comprehension stage and the third level is the *proposed assessment* for the projection stage.

In order to evaluate the ranking capability for the in-line situational assessment demonstrated in 4.2.3, the evaluation process is divided into two phases.

The first phase substitutes relevant values for the situational assessment components such as the *Ground Truth* and *Proposed Assessment*. Since we have already evaluated the underlying situational assessment in section 4.2.4, the situational assessment components for the multi projection environment are defined in sections 4.2.5 and 4.2.6. Therefore, the evaluation process presents the second phase for underlying evaluation.

The second phase is to quantify the ranking capability for a multi projection environment using the *Ranking Capability Score'* (RCS'). The quantitative assessment aims to evaluate the proposed assessment in comparison to the ground truth for each projection technique.

Firstly, the (RCS') quantifies the first projection technique, where the perception stage has identified two AoI. This is done by evaluating the proposed assessment (shown in Table 4.4)in comparison to the absolute truth of the situation (provided in Table 4.1).

Secondly, it quantifies the second projection technique, where the comprehension stage has proposed three AoI. This is done by evaluating the proposed assessment (shown in Table 4.5) in comparison to the absolute truth of the situation (provided in Table 4.2).

Thirdly, the enhanced proposes metric quantifies the third projection technique, where the projection stage has anticipated four AoI. This is done by evaluating the proposed assessment (shown in Table 4.6), in comparison to the absolute truth of the situation (provided in Table 4.3).

4.3.3 Quantification Process

This section evaluates the ranking capability for a multi projection environment, using the the *Ranking Capability Score'* (RCS'). It begins by quantifying the ranking

instances for the perception and comprehension stage. Finally, the evaluation process quantifies the ranking instances for the projection stage.

Information Perception

In the previous section 4.2.7 we substituted four values in relation to the perception stage. The first one is the current number of hops CoH_i (defined in equation (3.11). Secondly, we substituted the actual number of hops AoH_i (defined in equation (3.10) followed by finding the difference in number of hops DoH_j (in equation 3.12). Finally, the reference number of hops RoH_i (defined in equation 3.13) was determined.

Now, the underlying process substitutes three values for the Ranking Capability Score' (RCS'). The first is the reversed set for the actual number of hops AoH'_j. (defined in equation 4.1). It is substituted as follows: $AoH'_j = \{AoH'_1(0), AoH'_2(1)\}$.

Secondly, it computes the worst state for the difference in number of hops DoH'_j. (defined in equation 4.2). It is substituted as follows: $DoH'_j = \{DoH'_1(1), DoH'_2(0)\}$.

Thirdly, it computes the worst state for the reference in number of hops RoH'_j (defined in equation 4.3). It is substituted as follows: $RoH'_j = \{RoH'_1(2), RoH'_2(0)\ right\}$

Finally, it computes the Ranking Capability Score (RCS') for the information perception, comprehension and projection, (defined in equation 4.4) as follows:

$$RCS' = \frac{\left(\sum_{j=1}^{2} = roh_1(0) + roh_2(0+0)\right)}{\left(\sum_{j=1}^{2} = roh_1(2) + roh_2(2+0)\right)} = \frac{0}{4}.$$

Information Comprehension

In the previous section 4.2.7 we substituted four parameters in relation to the comprehension stage. The first is the current number of hops CoH_i (defined in equation 3.11) Secondly, we substituted the actual number of hops AoH_i (defined in equation 3.10) followed by determining the difference in number of hops DoH_j (defined in

equation 3.12) Finally, the reference number of hops RoH_i (defined in equation 3.13) was determined.

Now, we substitute another three parameters for the Ranking Capability Score (RCS'). The first is the reversed actual number of hops AoH'_j, (defined in equation 4.1); it is substituted as follows: $AoH'_j = \{AoH'_1(0), AoH'_2(1), AoH'_3(2)\}$

Secondly, we compute the worst state for the difference in number of hops DoH'_j which is defined in equation 4.2.
It is substituted as follows: $DoH'_j = \{DoH'_1(2), DoH'_2(2), DoH'_3(0)\}$.

Thirdly, we compute the worst state for the reference in number of hops RoH'_j which is defined in equation 4.3. It is substituted as follows: $RoH'_j =$
$\{RoH'_1(6), RoH'_2(4), RoH'_3(0)\}$.

Finally, it computes the *Ranking Capability Score'* (RCS') for the comprehension stage (defined in equation 4.4) as follows:

$$RCS' = \frac{\left(\sum_{j=1}^{3} = roh_1(6) + roh_2(6+4) + roh_3(6+4+0)\right)}{\left(\sum_{j=1}^{3} = roh_1(6) + roh_2(6+4) + roh_3(6+4+0)\right)} = \frac{26}{26}.$$

Information Projection

In the previous section 4.2.7 we substituted four parameters in relation to the projection stage. The first was the current number of hops CoH_i (defined in equation (3.11). Secondly, we substituted the actual number of hops $AoH_i i$ (defined in equation (3.10) followed by the difference in number of hops DoH_j (defined in equation 3.12). Finally, the reference number of hops RoH_i (defined in equation 3.13) was determined.

Now, we substitute another three parameters for Ranking Capability Score (RCS'). The first is the reversed actual number of hops AoH'_j, (defined in equation 4.1). It is substituted as follows: $AoH'_j = \{AoH_1(0), AoH_2(1), AoH_3(2), AoH_4(3)\}$.

Secondly, we compute the worst state for the difference in number of hops DoH'_j, (defined in equation 4.2). It is substituted as follows: $DoH'_j = \{DoH'_1(3), DoH'_2(4), DoH'_3(3), DoH'_4(0)$.

Thirdly, we compute the worst state for the reference in number of hops RoH'_j, (defined in equation 4.3). It is substituted as follows: $RoH'_j = \{RoH'_1(12), RoH'_2(12), RoH'_3(6), RoH'_4(0)\ right\}$.

Finally, we compute the *RankingCapabilityScore* (RCS') for the perception stage (defined in equation 4.4) as follows:

$$RCS' = \frac{\left(\sum_{j=1}^{4}{=roh_1(0)+roh_2(0+6)+roh_3(0+6+4)+roh_4(0+6+4+0)}\right)}{\left(\sum_{j=1}^{4}{=roh_1(12)+roh_2(12+12)+roh_3(12+12+6)+roh_4(12+12+6+0)}\right)} = \frac{26}{96}$$

4.3.4 Assessment Results

This section presents comparative results for the *RankingCapabilityScore* (RCS'), versus, the inline situational assessment demonstrated in 4.2.3. The discussion encompasses two levels of evaluations. At first glance, Figure 4.8 part (a), (b) and (c) presented the *Ranking Capability Score*$'$ (RCS'), against the information perception, comprehension and projection . However, at a second glance, Figure 4.8 part (d) demonstrates a comparative between the *RankingCapabilityScore* (RCS'), versus three qualitative states.

The first level has three points of comparisons against the multi projection environment. The first is shown in Figure 4.8-(a) where the first projection technique scores (0). This means the perception stage has ranked the identified threats (shown in Table 4.4) as perfect as the ground truth (shown in Table 4.1). Hence, the scoring scheme for the *Ranking Capability Score*$'$ (RCS') in relation to the number of identified threats is shown between (0) and (1). From these, (0) presented the good state, whilst (1) represents the worst state for the underlying situational assessment.

The second point of comparison is shown in Figure 4.8 part (b). The second projection technique scores (1). This means, the comprehension stage has ranked the proposed threats (shown in Table 4.5) as opposed to the ground truth (shown in Table 4.2). Hence, the scoring scheme for the *RankingCapabilityScore* (RCS') in relation to the number of proposed threats is shown between (0) and (1). From which, (0) presented the good state, whilst, (1) represented the worst state for the underlying situational assessment.

The third point of comparison is shown in Figure 4.8 part (c). The third projection technique scores (0.27). This means, the projection stage has ranked the anticipated threats (shown in Table 4.6) as opposed to the ground truth (shown in Table 4.3). Hence, the scoring scheme of the *RankingCapabilityScore* (RCS') ,in relation to the number of anticipated threats, is shown between (0) and (1). From these, (0) represented the good state, whilst (1) represented the worst state for the underlying situational assessment.

Similarly, Figure 4.8, part (d) presents the ranking capability score versus the three qualitative states. The *Good State* scores (0) for the first projection technique in which the perception stage has ranked two intrusion activities as perfect as the ground truth. The *Degraded State* scores (0.13) for the second projection technique in which, the projection stage has ranked four anticipated threats as not perfect as the ground truth. The *Bad State* scores (1) for the third projection technique where the comprehension stage has ranked three intrusion activities as opposed to the ground truth.

In the light of the three qualitative states, the *Degraded State* for the projection stage scores (0). The *Bad State* for the comprehension stage scores (1). The *degraded state* for the projection stage scores (0.27). *RankingCapabilityScore* (RCS') has quantified each projection technique appropriately. It has provided a reliable scoring scheme for three different ranking instances regardless of the dynamic changes for

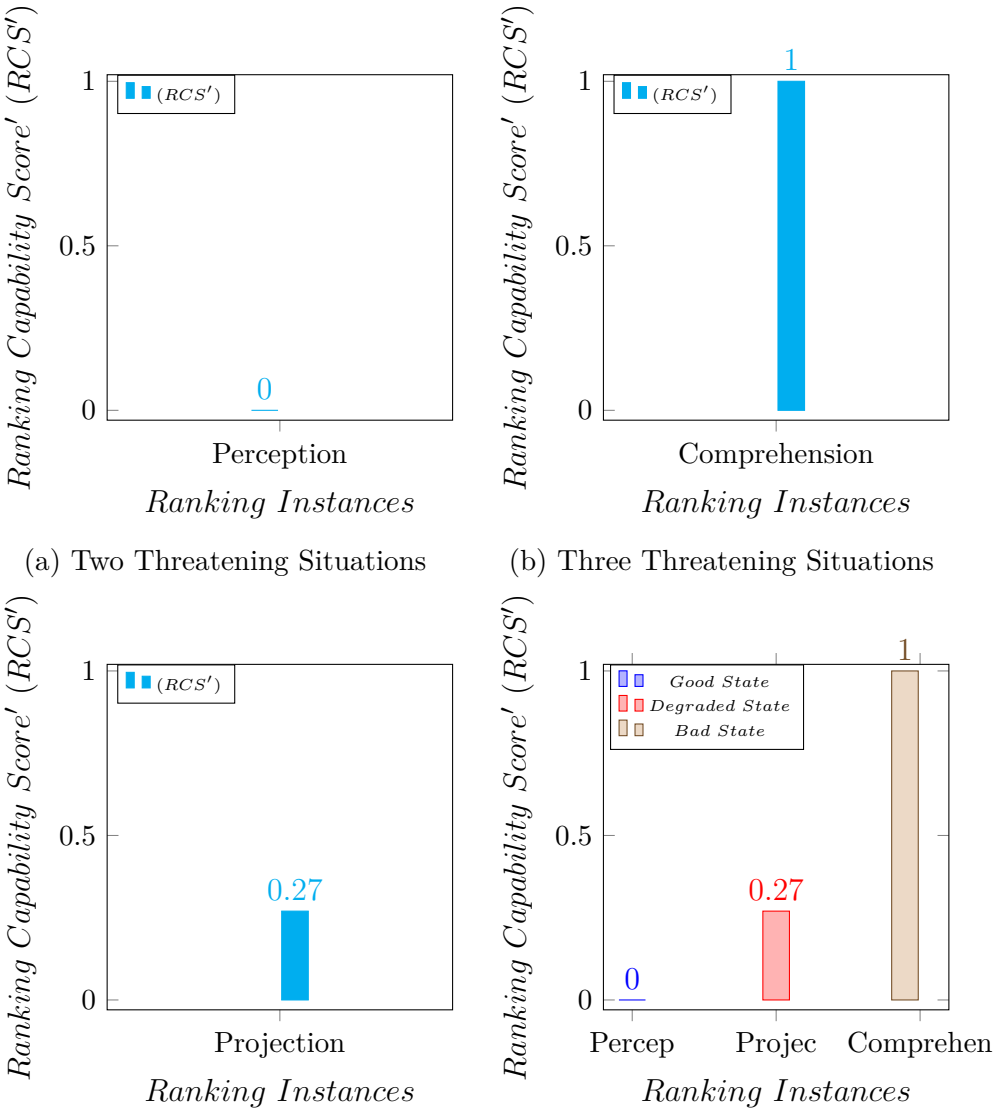


Fig. 4.8 Measuring the Ranking Capability for Multi Projection environment using the *Ranking Capability Score′* (RCS')

the number of emerging threats in a multi projection environment. this means the enhanced RCS has overcome the knowledge representation problem concerning with the user system relation.

This section examined the *Ranking Capability Score′* (RCS') with the help of a case study. However, the demonstrated situational assessment showed only the capability of the proposed metric against two or three ranking instances. The next

section conducts a comparative evaluation for the *Ranking Capability* ($Score'$) versus the *Ranking Capability Score* (RCS) with two levels of assessments. The first level is a quality based evaluation used to examine the scoring scheme against all the ranking instances for three different scenarios. The first scenario, where the real time system is reporting two tracking activity for the decision making resources, second scenario versus three tracking activity, and the third scenario versus four tracking activity. The second level is a reliability based evaluation used to examine the scoring scheme against the three qualitative states defined previously; the good state, degraded state and bad state.

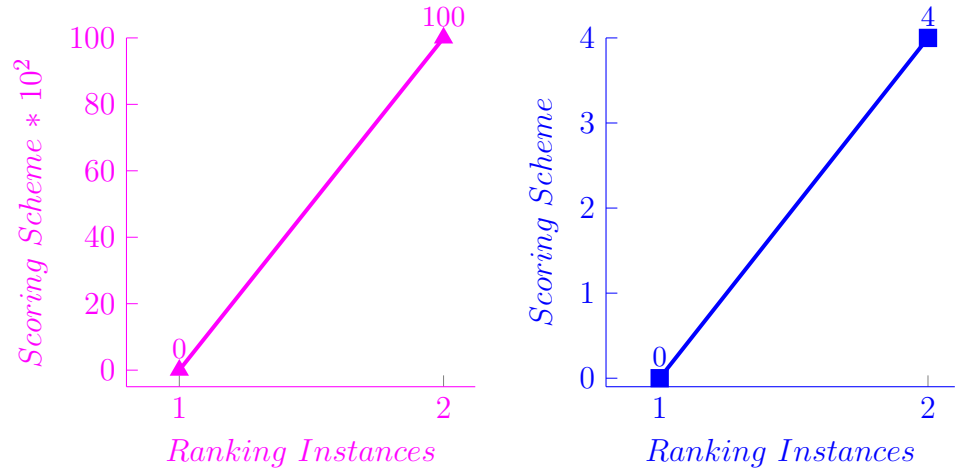
4.4 Comparative Evaluation

This section presents a comparative evaluation for the *Ranking Capability* ($Score'$) versus the enhanced Ranking Capability Score (RCS). The evaluation process encompasses three parts. The first part examines both scoring schemes, versus all the ranking instances over three separates scenarios. The second part presents a comparative result for both scoring schemes, versus the demonstrated case study in section 4.2.3. The third part illustrates a comparative result based on the reliability of both scoring schemes, versus the three qualitative states likely to occur during a real time operation.

4.4.1 Quality Based Evaluation

This section presents a comparative evaluation for the (RCS') versus the (RCS). This is to examine both scoring schemes against their intended purpose. Each performance metric is expected to provide unique scores against all ranking instances over three different scenarios. Each of these scenarios encompassed different numbers of prioritised events. Furthermore, in the first scenario, we have assumed the proposed assessment

output has identified two prioritised events. In the second scenario, the SWA system has proposed three prioritised events. In the third scenario, the proposed assessments outputs has anticipated four prioritised events.



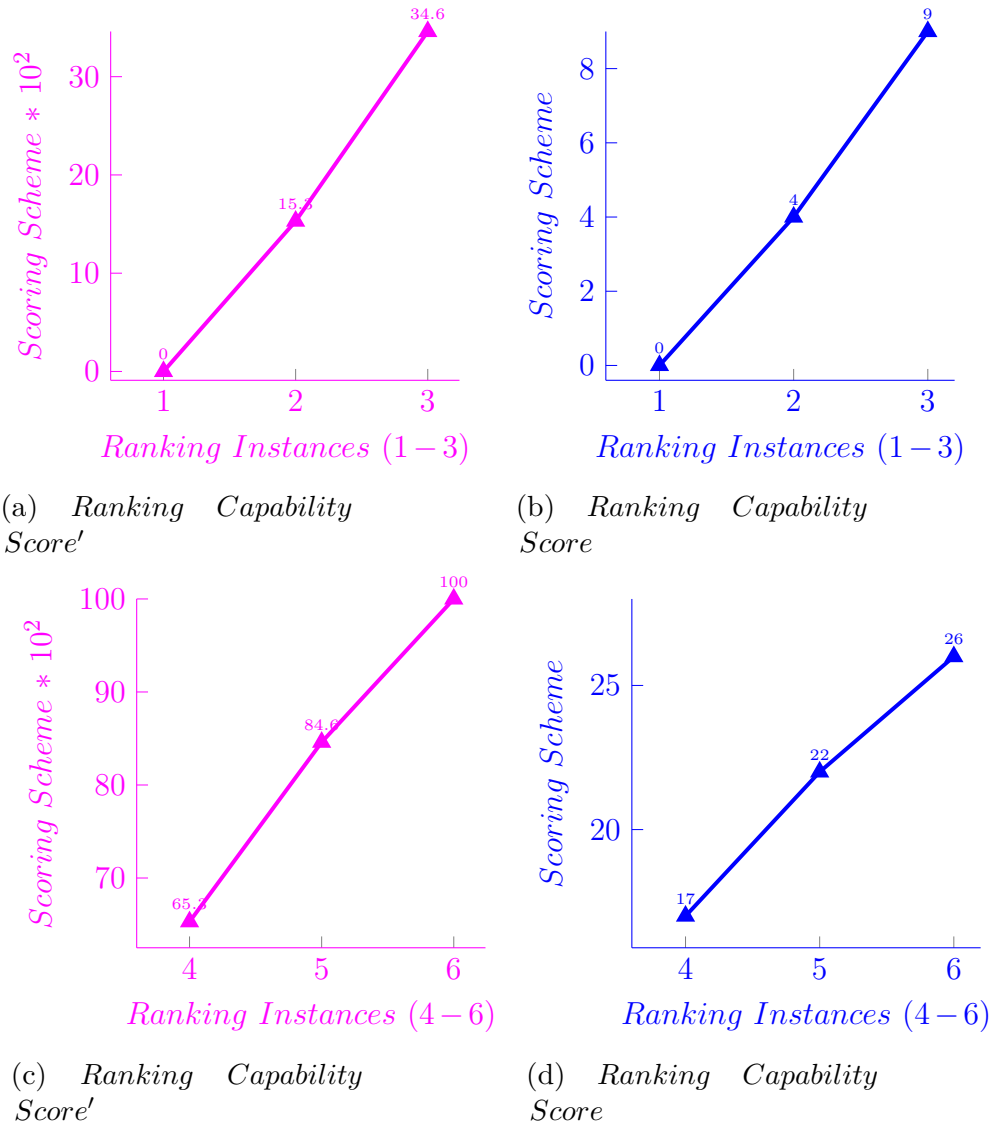
(a) *Ranking Capability Score′*

(b) *Ranking Capability Score*

Fig. 4.9 Quality Based Evaluation for RCS versus $RCS\prime$. $(2!) = 2$ *Ranking Instances*

Originally, the RCS' and RCS have been designed to evaluate the ranking capability in relation to the prioritisation process. This was under a scenario where the SWA system was reporting only important activities, but with different severity paradigms. This section examines both performance metrics against their intended purpose; each metric is expected to provide a unique score for all ranking instances against the three different scenarios. Initially, the underlying process examines both scoring schemes against the first scenario, where the SWA system is reporting two threatening activities (results are shown in Figure 4.9) and then against three threatening situations (results are shown in Figure 4.10) and finally, against four anticipated threats (results are shown in Figure 4.11).

The RCS' provides unique scores for all ranking instances in three different scenarios, as shown in Figures (4.9), (4.10) and (4.11). In the first scenario, the first ranking instances score 0, indicating, that the SWA system has ranked the emerging situations



(a) *Ranking Capability Score′*

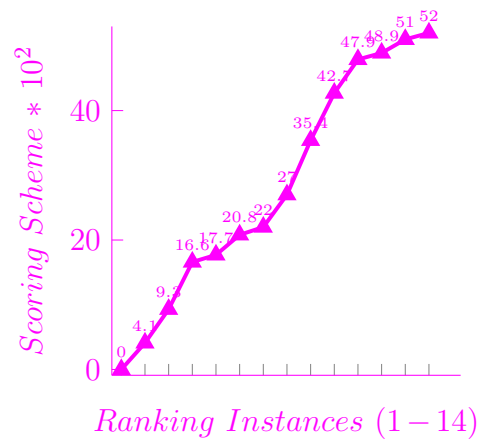
(b) *Ranking Capability Score*

(c) *Ranking Capability Score′*

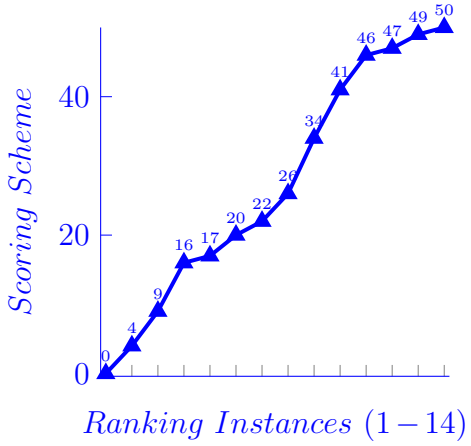
(d) *Ranking Capability Score*

Fig. 4.10 Quality Based Evaluation for RCS versus RCS'. $(3!) = 6$ *Ranking Instances*

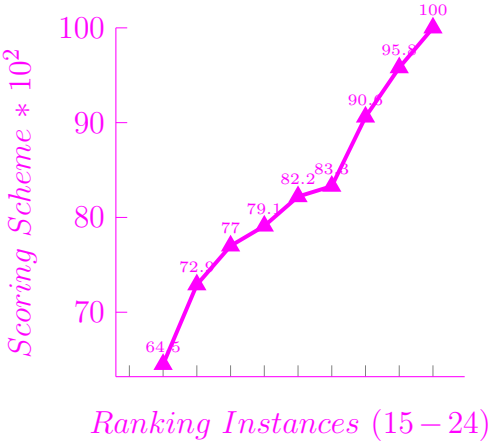
as perfect as the ground truth, while the second ranking instances scores (1) indicating that the SWA system has ranked the emerging threats as opposed to the ground truth. The second and third scenarios are shown in 4.10 and 4.11. The RCS' provide unique scores against all the ranking instances over the underlying situational assessment. This is because it has been designed to evaluate the underlying situational assessment, specifically under a scenario where the SWA system is reporting only distinct threatening situations but with different severity paradigms.



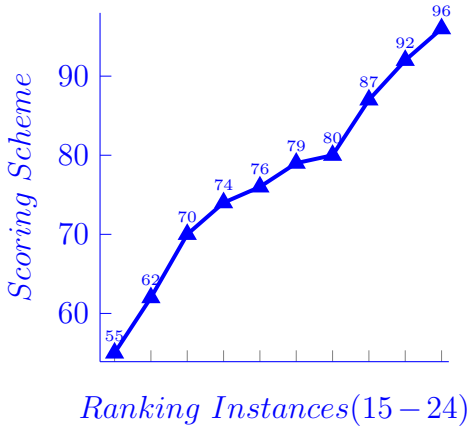
(a) *Ranking Capability Score′ RCS′*



(b) *Ranking Capability Score*



(c) *Ranking Capability Score′ RCS′*



(d) *Ranking Capability Score*

Fig. 4.11 Quality Based Evaluation for RCS Versus RCS'. $(4!) = 24$ *Ranking Instances*

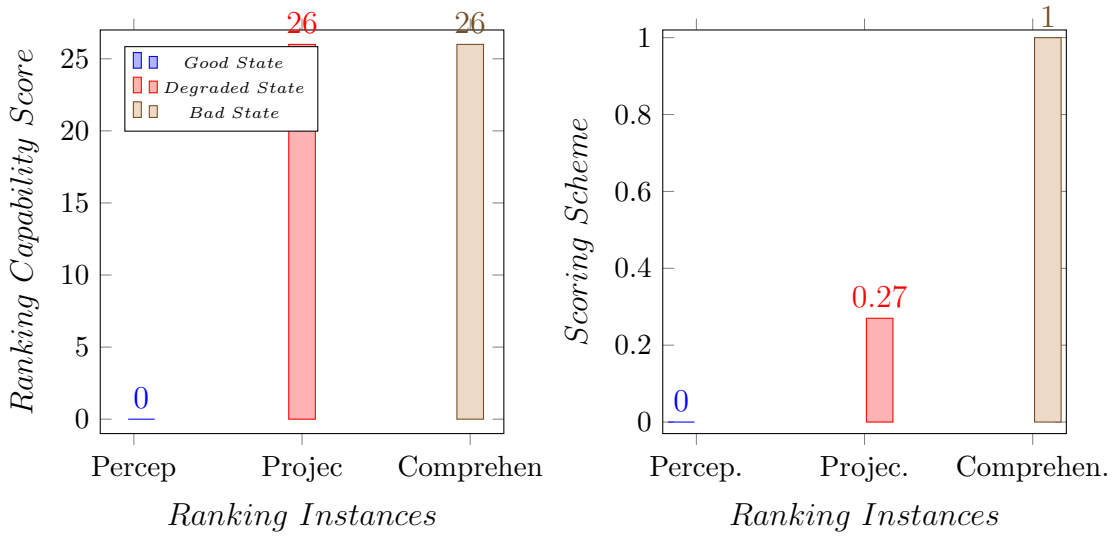
The RCS' provides unique scores for all ranking instances in three different scenarios, as shown in Figures (4.9), (4.10), (4.11). In the first scenario, the first ranking instances score 0, indicating that the SWA system has ranked the emerging situations as perfect as the ground truth, while the second ranking instances scores (2), indicating that the SWA system has ranked the emerging threats as opposed to the ground truth. The second and third scenarios are shown in Figures 4.10 and 4.11. The RCS provides unique scores for all the ranking instances over the underlying situational assessment. That

is because it has been designed to evaluate the ranking capability for the underlying scenario. Specifically, under situational assessment where the SWA system reports only important activities but with different severity paradigms.

This section showed that both scoring schemes successfully quantified the prioritisation process over three separate scenarios. The next section further examines both scoring schemes concerning the knowledge representation problem for the demonstrated case study in section 4.2.3.

4.4.2 Case Study Based Evaluation

This section presents a comparative evaluation for the RCS' and the RCS. This is to examine both scoring schemes against the multi projection environment. In other words, each performance metric is expected to provide unique scores against all ranking instances for the demonstrated case study 4.2.3, regardless of the dynamic changes for the number of emerging threats.



(a) *Ranking Capability Score* (b) *Ranking Capability Score* RCS'

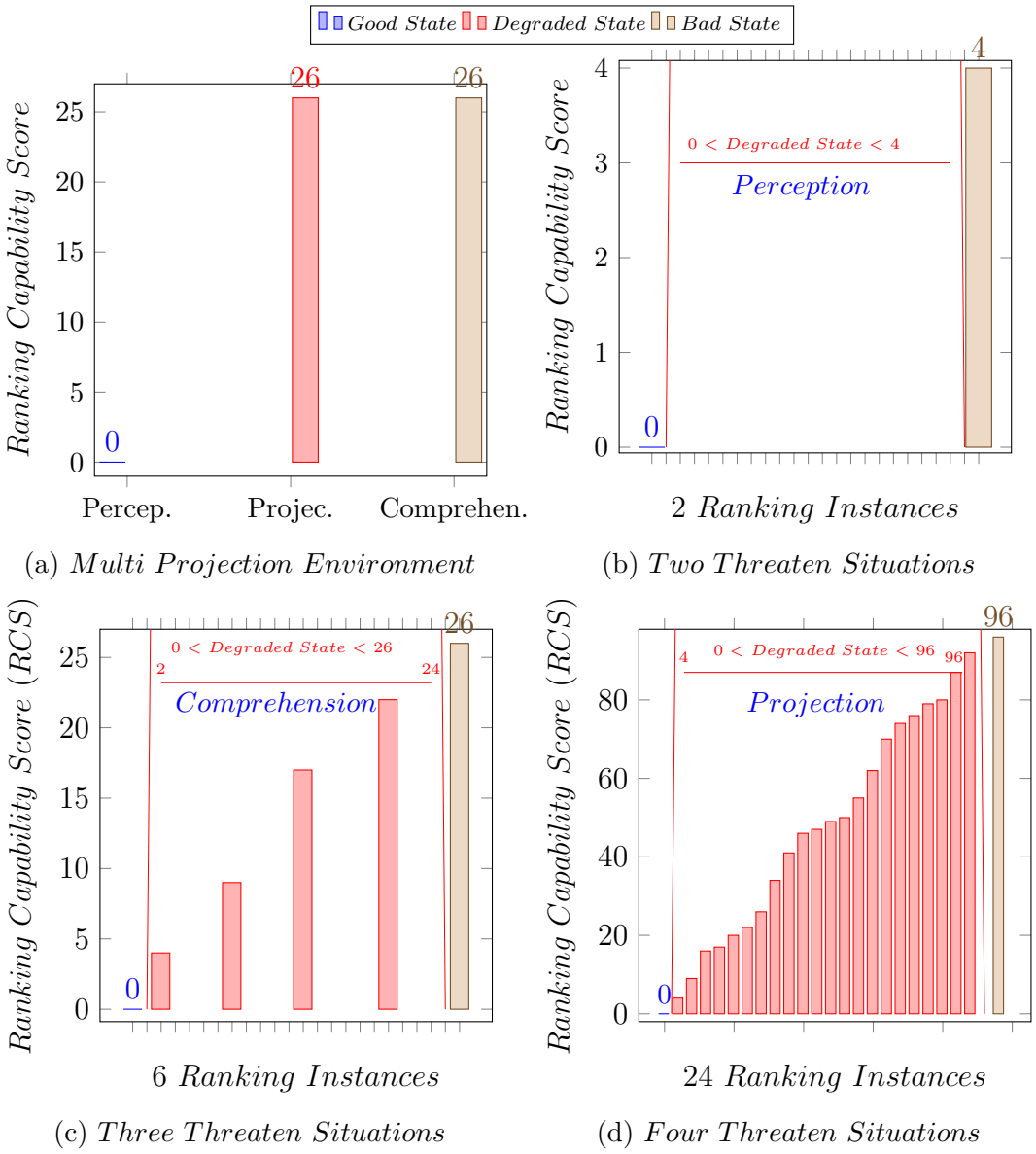
Fig. 4.12 Measuring the Ranking Capability for Multi Projection Environment

According to the demonstrated case study in section (4.2.3), the first projection technique (perception stage) has ranked the identified threats as perfect as the ground truth, while the second projection technique (comprehension stage) has ranked the proposed threats as opposed to the ground truth. However, the third projection technique has ranked the anticipated threats as not perfect as the ground truth. In regards to this, the evaluation process aimed to examine ranking capabilities for each projection technique, respectively. In other words, each performance metric is expected to generate unique scores for different ranking instances. Figure 4.12 presented the capability score against the multi projection environments.

Initially, the RCS in Figure 4.12-(a) scores (0) for the perception stage and then it scores 26 for the projection and comprehension stages. However, each projection technique has different ranking paradigms for the identified situations. This means the underlying performance metric has not provided a reliable score for different ranking instances concerning the demonstrated case study.

Similarly, the RCS' in Figure 4.12-(b) scores (0) for the perception stage indicating that the first projection techniques ranked the emerging threats as perfect as the ground truth. It then scores 0.13 for the projection technique, indicating that the projection stage has ranked the anticipated threats as not perfect as the ground truth. Finally, it scores (1) for the third projection technique. This means, the comprehension stage has ranked the emerging situation as opposed to the ground truth. Indeed, based on the obtained results, the ranking capability score$'$$has quantified the ranking capability for the underlying scenario appropriately.$

Furthermore, the evaluation process has conducted further assessment for the RCS in Figure (4.13) and for RCS' in Figure (4.14) to find an answer for the underlying results.

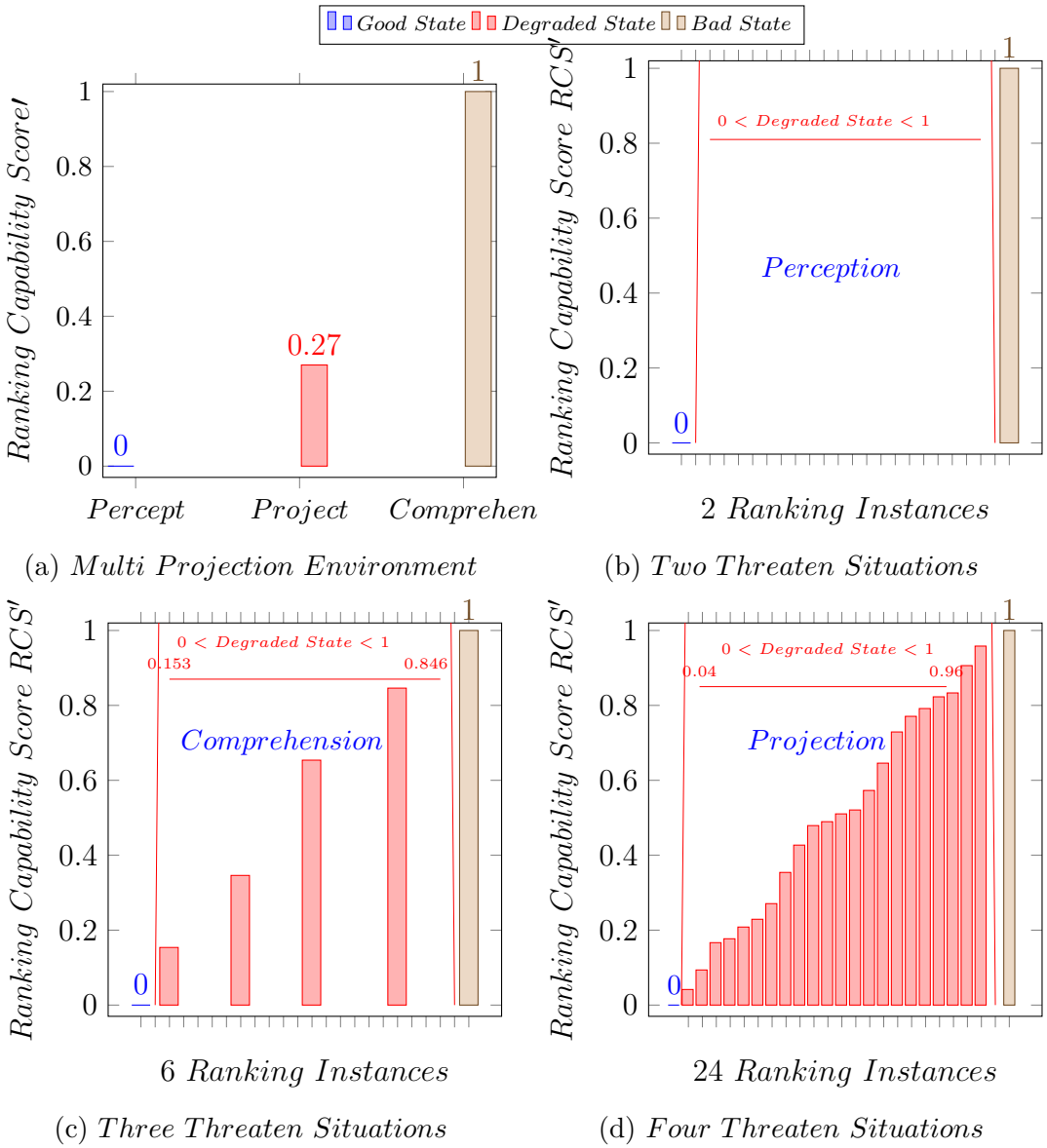


(a) *Multi Projection Environment* (b) *Two Threaten Situations*

(c) *Three Threaten Situations* (d) *Four Threaten Situations*

Fig. 4.13 Ranking Capability Score (RCS) versus Number of Identified Threats

The *RCS* in Figure (4.13) scores (0) for the good state in relation to number of identified threats(shown in part (*a*)). However, it scores 26 for the degraded state, in relation to the number of emerging threats showing in part (d). Furthermore, it also scores 26 for the bad state, in relation to the number of emerging threats (shown in part (c)). Practically, from the quantitative assessment perspective, the *RCS* has quantified each projection appropriately. However, from the qualitative perspective,



(a) *Multi Projection Environment* (b) *Two Threaten Situations*

(c) *Three Threaten Situations* (d) *Four Threaten Situations*

Fig. 4.14 *Ranking Capability Score* RCS' Versus Number of Identified Threats

the underlying metric does not provide a reliable scoring scheme for multi projection environments, specifically under a situational assessment where the number of emerging threats are dynamically changing.

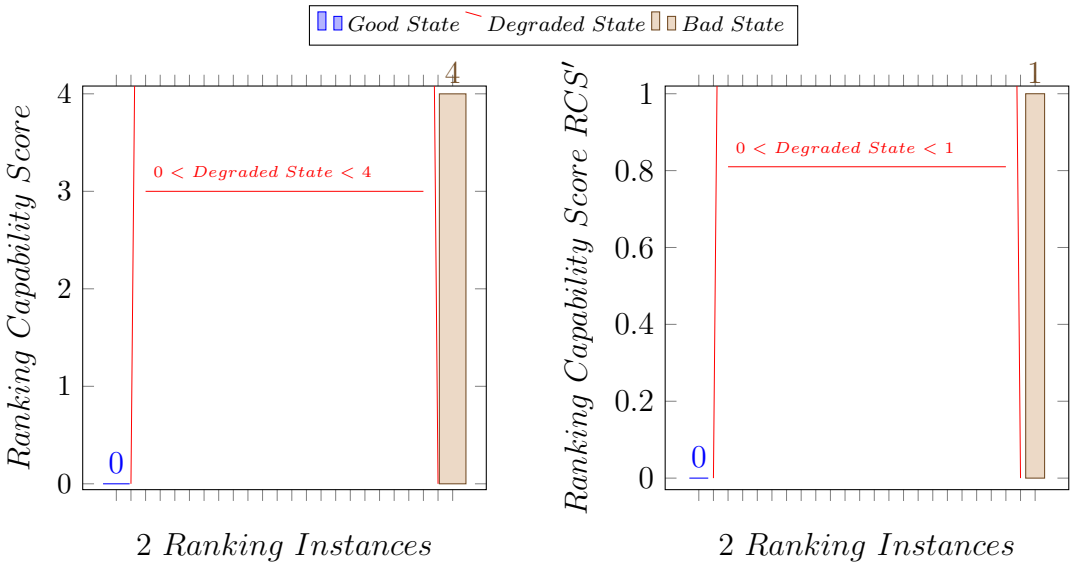
The RCS' in Figure (4.14) scores (0) for the good state in relation to number of identified threats(shown in part (a)). However, it scores 0.13 for the degraded state, in relation to the number of emerging threats showing in part (d). Furthermore, it

scores 1 for the bad state ,in relation to number of emerging threats (shown in part (c). Practically, from both quantitative and qualitative assessment perspectives, the RCS has quantified each projection appropriately. Therefore, it provides a reliable scoring scheme for multi projection environments regardless of the dynamic changes for the number of emerging threats in real time operation.

The next section further examines both scoring schemes against the qualitative states in three different scenarios.

4.4.3 Reliability Based Evaluation

This section presents a comparative evaluation for the RCS' and the (RCS). This is to examine both scoring schemes against three qualitative states. Each performance metric is expected to provide appropriate scores against different states regardless of the dynamic changes for the number of emerging threats.

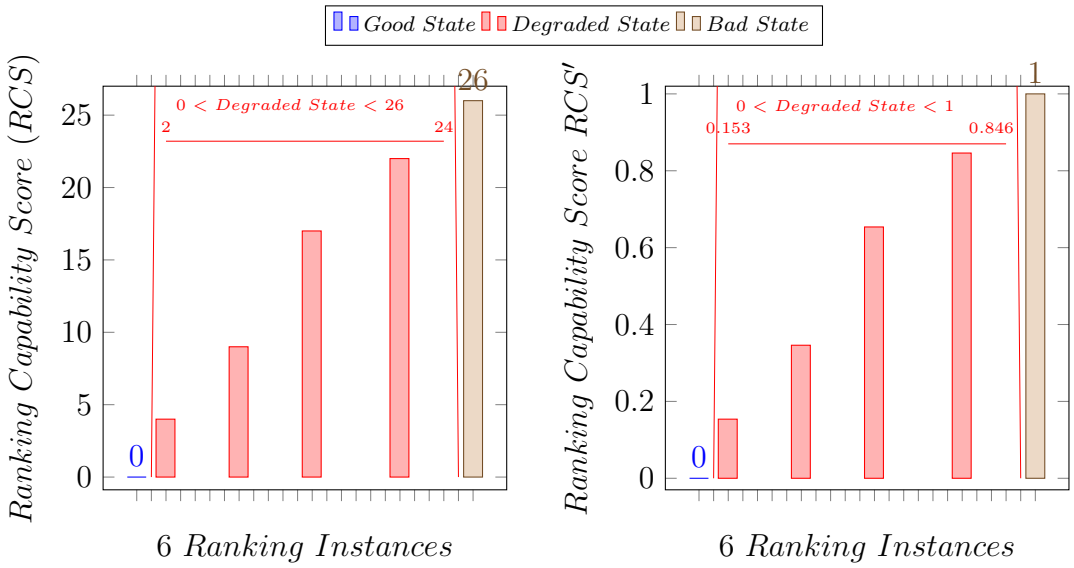


(a) *Ranking Capability Score* (b) *Ranking Capability Score RCS'*

Fig. 4.15 Qualitative States versus Scoring Scheme. (2!) 2 Ranking Instances

The (RCS) in Figure (4.15- a) scores (0) for the *Good State*, and (4) for the *Bad State*. However, when the number of emerging threats changes to three AoIs, as shown

in Figure (4.16- a), the RCS scores (0) for the *Good State* and (26) for the *Bad State*. Likewise, the *Degraded State* scores (2) − (24). Lastly, when the number of emerging threats increased up to four emerging threats, as shown in Figure (4.17-a), it scores (0) for the *Good State* and (96) for the *Bad State*. It scores between (4) − (96) for *Degraded State*.

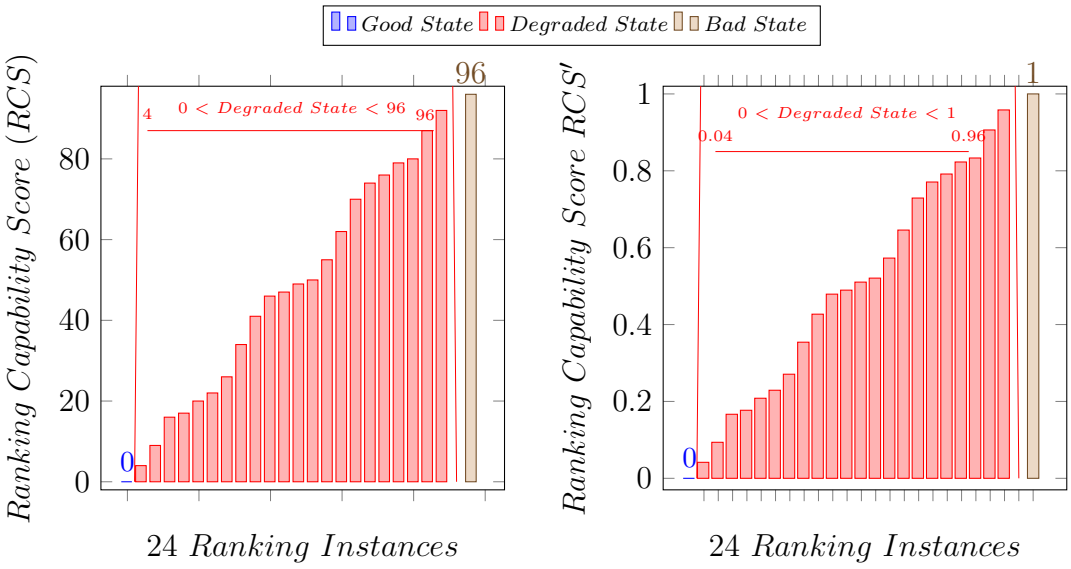


(a) *Ranking Capability Score* (b) *Ranking Capability Score* RCS'

Fig. 4.16 Qualitative States versus Scoring Scheme. (3!) 6 Ranking Instances

The RCS' in Figure (4.15-b) scores (0) for the *Good State*, and (1) for the *Bad State*. However, when the number of emerging threats changes to three AoI, as shown in figure (4.16-b), the RCS' scores (0) for the *Good State*. It scores between (0.1538) to (0.8461) for the *Degraded State* and it scores (1) for the *Bad State*. Lastly, when the number of emerging threats changes to four AoI, as shown in (4.17-b), it scores (0) for the *Good State* and (1) for the *Bad State* while it scores between (0.0416)to(0.9583) for the *Degraded State*,

The *Ranking Capability Score* RCS' provides a reliable scoring scheme for the three qualitative states, regardless of the number of emerging threats. It scores (0) for the good states and (1) for the bad state, in addition to that, it provides an appropriate



(a) *Ranking Capability Score* (b) *Ranking Capability Score RCS′*

Fig. 4.17 Qualitative States versus Scoring Scheme. (4!) 24 Ranking

scoring scheme for the degraded state between $[0-1]$. However, the RCS provides an inappropriate scoring scheme for the three different states. The scoring scheme for the bad and degraded state changed in relation to the number of emerging threats. This can cause uncertainty in the decision making resource during a real time operation, where the number of emerging threats are dynamically changing.

4.5 Conclusion

This chapter has examine the knowledge representation of the Ranking Capability Score (RCS) against three distinct qualitative states (demonstrated in section 4.2.1). The evaluation process examined the proposed scoring scheme, specifically under a contextual scenario and when the real time system is conducting three levels of assessments against the identified situation. However, each level is reporting different numbers of tracking activities for the analysis team.

The result obtained in Figures (4.1), (4.2), (4.3) showed that the RCS for the three qualitative state changed when the proposed number of tracking activities changed respectively. This means, the underlying performance metric does not provide reliable scores for the decision making resources, specifically during the time when the real time system has been configured to perform multi perspective measurements against the emerging situation (as has been demonstrated in section 4.2.8).

In regards to this, we introduced an enhanced method for the RCS in section 4.3 to provide a reliable score for addressing the knowledge representation problem concerning the user refinement (level 5) of the JDL, especially as the real time system is configured to assess any dynamically monitored environment with multilevel of assessments.

Furthermore, we conducted three levels of assessment to validate the enhanced scoring scheme against its intended purpose. The first level of assessment was a case study based evaluation (showed in section 4.2.2). The demonstrated evaluation showed only the capability of the proposed performance metrics against two or three ranking instances (as has been shown in Figure 4.8).

With this in mind, we conducted a quality based evaluation. The evaluation process aimed to examine the scoring scheme against its intended purpose. The proposed performance metric provided unique scores against different ranking instances concerning the prioritisation process of a real time system. The obtained results showed that the RCS' provided appropriate scores against all the ranking instances over three separate scenarios.

Moreover, section 4.4 conducted a comparative evaluation. The evaluation process encompassed three levels of assessments to illustrate the differences between the RCS' scoring scheme versus the RCS scoring scheme. The first level was a quality based evaluation (demonstrated in 4.4.1). The second level was a case study based evaluation

(demonstrated in 4.4.2). The third level was a reliability based evaluation (demonstrated in 4.4.3).

The obtained results for the quality based evaluation showed that the initial method of the RCS provided unique scores against all the ranking instance over three different scenarios. Additionally, it provided an inappropriate scoring scheme for the case study based evaluation (shown in Figure 4.12-(a)). Moreover, it has provided unreliable scores for the three different qualitative states, as has been demonstrated in Figures (4.15-(a)), (4.16-(a)) and (4.17-(a)).

The enhanced method of the RCS' provided unique scores against all the ranking instance over three different scenarios (showed in Figures (4.9-b),(4.10-(a),(c)) and (4.11-(a),(c)). It also provided appropriate scores against the case study based evaluation (shown in Figure 4.12-b). Finally, it provided a reliable scoring scheme for the analysts team, as has been demonstrated in Figure (4.15-(b)), (4.16-(b)) and (4.17-(b)).

Indeed, the *Ranking Capability Score'* has provided a better scoring scheme in comparison to the previous metric, in term of providing a representative scores for the user perception concerning the three qualitative states mention above, in particular under an extended scenario where the number of prioritised events are dynamically filtered during a real time operation.

Chapter 5

Scheduling Capability Score (SCS)

5.1 Introduction

The data fusion community has proposed multiple levels of situational assessment[26] to enable timely responses during real time operations[19]. During assessment the situational awareness system can report different classes of tracking activities; each class has a different degree of importance. Ideally, the decision-making team are interested in viewing the most important class. However, the real time system does not always schedule the identified classes in respect of their degree of importance. This can be due either to a lack of knowledge about the dynamically monitored environment or to a configuration problem in the aggregation, classification or correlation techniques of the situational awareness system.

The first contribution of this chapter is introducing the Scheduling Capability Score (SCS) for evaluating the ranking capability of a situational awareness (SWA) system. The proposed performance metrics have been designed and evaluated using an analytic approach. The modelling scheme represents the SWA system outputs mathematically, in the form of a list of activities. Such methods allowed the evaluation process to

conduct a rigorous analysis of the scheduling process, despite any constraint related to a domain-specific configuration.

Furthermore, this work conducts a comparative evaluation with existing performance metric[74]; the evaluation methodology encompassed two levels of assessments with the first level being a case study based evaluation. The second level is a quality-based evaluation to examine the underlying metrics against their intended purpose.

The second contribution of this work is to deliver a method to analyse the computational complexities involved in evaluating the prioritisation and scheduling processes for two distinct operations. The first operation is during the assessment stage where the evaluation process required computes only the necessary values to assess the ranking capability of the real time system. The second operation is during the optimisation stage; where the evaluation process is required to compute more values to assess all the ranking instances for any given scenario.

The chapter is divided as follows: the first section illustrates a guidance case study in which the ranking capability of a real-time system is affected due to a configuration problem during the aggregation stage for the multiple sensor information fusion. The second section introduces a modelling scheme for representing the ranking capability problem of a real time system.

The third section presents the developmental phase for the proposed performance metric. The fourth section demonstrates two levels of evaluation to examine the proposed scoring schemes against their intended purpose.

The fifth section discusses the computational complexity issues for two different operations involved in evaluating the prioritisation and scheduling processes for a real time system. The first operation is during the assessment stage where the underlying performance metric is required to compute only essential values. The second operation occurs when the potential performance metric is required to compute more values to

assess the optimisation technique concerning the ranking capability of the situational awareness system. Finally, we will discuss our findings and future work.

5.2 Ranking Capability Issue in Real-Time System

This section presents a guidance case study, from the cyberspace domain [74], in which the real time system is experiencing ranking capability issues during multilevel situational assessments due to a lack of configuration. The second section demonstrates how the real time system is reporting the identified situation.

5.2.1 Overviews

According to the underlying scenario adapted from [74], the inline situational assessments model has been configured to represent the identified situation in a list of prioritised events. However, during the aggregation process of multiple information, the SWA system has identified two classes of events.

- The first is a complex event which has high impact against the protected environment, namely *The Activity of Interest*.
- The second is a complex event regarded as a normal activity, which has minimal or no impact against the dynamical environment.

The real-time system has performed two levels of assessments against the identified situation but the system reported undesirable events due to a configuration problem in the aggregation process. Therefore, the perception and comprehension stages have not ranked the identified events as perfect as the ground truth.

The next section demonstrates how the real-time system reports the identified situation, during the time when the system was experiencing ranking capability issues.

5.2.2 Reporting Different Classes of Prioritised Events

Following the contextual scenario in the previous section, the perception stage has reported the identified activities at the time of their occurrence without any contextual order. The proposed assessment is shown in Table 5.1. However, after conducting further assessments against the identified situation, the SWA system, at the comprehension stage, has ranked the activities of interest into a contextual order; this is shown in Table 5.2.

Apparently, the ranking instance at the comprehension stage is better than the perception stage, in terms of shifting the $AOIs$ over the normal events. Nevertheless, the underlying ranking paradigms are not as perfect as the ground truth. This is shown in Table 5.3.

Table 5.1 Proposed assessment at the perception stage adapted from[74]

Proposed Assessment	**Activity**	**Priority**
PA_0	Activity	4
PA_1	Activity	3
PA_2	Activity (AoI)	2
PA_3	Fragmented Activity	-
PA_4	Activity	5
PA_5	Activity (AoI)	1
PA_6	Activity not part of G.T	-
PA_7	Activity	6

G.T Ground Truth

The next section introduces the modelling scheme for the underlying case study in order to evaluate the scheduling process of the real time system.

Table 5.2 Proposed assessment at the comprehension stage adapted from[74]

Proposed Assessment	**Activity**	**Priority**
PA_0	Activity	4
PA_1	Activity (AoI)	1
PA_2	Activity (AoI)	2
PA_3	Activity	3
PA_4	Fragmented Activity	-
PA_5	Activity	5
PA_6	Activity not part of G.T	-
PA_7	Activity	6

G.T Ground Truth

Table 5.3 Identified activity at the ground truth adapted from[74]

Ground Truth	**Activity**	**Priority**
GT_0	Activity (AoI)	1
GT_1	Activity (AoI)	2
GT_2	Activity	3
GT_3	Activity	4
GT_4	Activity	5
GT_5	Activity	6
GT_6	Activity	7

5.3 Modelling Scheme

The first section presents the developmental phase of the situational assessments object. The second section presents the ground truth object and finally, the third section presents the proposed assessments outputs during real time operation.

5.3.1 Situational Assessments

According to the demonstrated case study, the situational assessment identified two classes of activity. The first is a number of complex events, with different priorities,

having significant impact against the monitored environment. This is defined in equation 5.1;

$$AoI_{\beta} = \{1, 2, \ldots \eta\} \tag{5.1}$$

where $1 \leq \beta \leq \eta$, $AoI_{\beta} = \beta$ And $\eta =$ The total number of important activities

The AoIs represent a set of complex events classified as important activities with high impact against the protected environment, while (β) represents the respective priority of each one. Therefore, the first activity AoI_1 found on the ground truth was prioritised with the value of (1). The next immediate activity was prioritised with the value of (2) and the last important activity found on the ground truth was prioritised with the value of (η), wherein, η is the total number of important activities concerning the emerging situation.

The second objects are complex events with minimal or no impact on the monitored environment. However, they have different priorities; these are defined in equation 5.2;

$$A_{\delta} = \{\eta + 1, \ldots \Psi\} \tag{5.2}$$

where $1 \leq \delta \leq \zeta$, $\Psi = \eta + \zeta$ And $\zeta =$ The total number of normal activities.

The A_{δ} represent a set of complex events classified as normal activities with minimal impact on the protected environment. Therefore, the first activity A_1 found on the ground truth was prioritised with the value of $A0I_{\eta} + 1$. The next immediate activity was prioritised with the value of $A_1 + 1$ and the last important activity found on the ground truth was prioritised with the value of $(A_{\zeta} + 1)$, wherein, ζ is the last normal activity found in relation to the identified situational assessment.

The third object of the situational assessment is called the Atomic Event (AE). This is defined below as:

$$AE_{\alpha} = \{AE_1, \ldots AE_{\sigma}\} \tag{5.3}$$

where $1 \leq \alpha \leq \sigma$ σ = The total number of atomic events $AE_{\alpha} = \alpha$

The AE_{α} represent a set of atomic events found unexpectedly among the proposed complex events in a real time operation. Usually, this occurs due to capability issues with the lower level of the SWA system, such as the correlation, aggregation and classification techniques.

Therefore, this class of event is at the bottom of the proposed list with no priorities. The activity of interest AoI_{β} has different degrees of importance; this is defined in equation 5.4:

$$DAoI_{\varsigma} = \{SOI_{N+1-\varsigma}\}, DAoI_1 > DAoI_2 > \ldots > DAoI_{\varsigma} \tag{5.4}$$

where $1 \leq \varsigma \leq \eta$ $N = total\ number\ of\ events\ being\ proposed\ by\ the\ situational\ assessment$

$$SoI_j = \{N - NoH_i\}, SoI_1 > SoI_2 > \ldots > SoI_N \tag{5.5}$$

where, $1 \leq j \leq N$, $1 \leq i \leq N$

$$NoH_i = N - i \tag{5.6}$$

where, $i = 1, \ldots, N$

The NoH_i indicates how many hops away from the bottom of the list each $i'th$ activity is. Conversely, it shows the predetermined order for each AoI with respect to their initial priority. For example, when considering a number of complex events occurring in the ground truth list, the least important activity should remain on the bottom of the list with 0 number of hops, while the immediate next important activity should have a number of hops equal to 1. Hence, the number of hops is incrementing by 1 up until the most important activities. The most important activity of all should have the maximum number of hops. This is called the total number of hopToH and is defined above in equation 5.6 as $ToH = N - 1$.

Ideally, during situational assessment in a real time operation, the most important activity found on the ground truth should have the highest score in terms of importance. Hence, the score of importance SoI_j is related to the severity of all activity being tracked by the SWA system; this is defined in equation 5.5.

$DAoI_\varsigma$ represents the score of importance for each activity in relation to their severity paradigms. For instance, the first activity of interest (AoI_1) $\ni$ severity scores $= DAoI_1$ with priority value of (1), while the second (AoI_2) $\ni$ severity score $= DAoI_2$ with priority value of (2) and the last AoI_β $\ni$ severity scores $= DAoI_\varsigma$ with priority value of η.

Furthermore, the normal class of tracking activity (A_δ) also has different degrees of importance regarding its severity paradigms; this is defined in equation 5.7 ;

$$DA_\Delta = \{SOI_{N+1-\Delta}\}, DA_1 > DA_2 > ... > DA_\varsigma \tag{5.7}$$

where $1 \leq \Delta \leq \zeta$ and N = *total number of events being proposed by the situational assessment.*

DA_Δ represents the score of importance of each activity in relation to their severity paradigms. For instance, the first normal activity A_1) $\ni$ severity scores $= DA_1$ with

priority value of $\eta+1$, while the second $(A_2) \ni$ severity scores $= DA_2$ with priority value of A_1+1 and the last $A_\beta \ni$ severity scores $= DA_\Delta$ with priority value of σ.

According to the underlying situational assessments, we have introduced different objects. Likewise, the number of hops NoH_i represents the predetermined order of all events being proposed by the situational assessment, and the SoI_j represents the score of importance concerning the severity paradigms of all events being tracked by the SWA system.

In the next step, the modelling scheme introduces the representations of the SWA system outputs, specifically, during the time where the real time system is reporting different classes of tracking activity in the real time operation. Therefore, the situational assessment (SA) contained a number of objects; this is defined in equation 5.8:

$$S.A_\gamma = \Big\{\Big\{AoI_\beta \ni DAoI_\varsigma\Big\} \prec \{A_\delta \ni DA_\Delta\} \prec \{AE_\alpha\}, \{NoH_i\}, \{SoI_j\}\Big\} \tag{5.8}$$

$$\text{where } 1 \leq j \leq N,\ 1 \leq \beta \leq \eta,\ 1 \leq \delta \leq \zeta,\ 1 \leq \alpha \leq \sigma,\ 1 < \gamma < N \mid \gamma = \eta+\zeta+\sigma$$

The next section introduces the ground truth component for the underlying situational assessment.

Ground Truth Outputs

The absolute truth about the situational assessment is defined in equation 5.9 as follows:

$$G.T_\iota = \Big\{\Big\{AoI_\beta \ni DAoI_\varsigma\Big\} \prec \{A_\delta \ni DA_\Delta\}\Big\} \tag{5.9}$$

$$\text{where } 1 \leq \beta \leq \eta,\ 1 \leq \delta \leq \zeta,\ 1 \leq \iota \leq \Psi \mid \Psi = \eta+\zeta$$

The next section introduces the proposed assessment component for the underlying scenario.

Proposed Assessment Outputs

We need to define how a real-time system might propose the identified situation during a real time operation; we do not know how different systems will rank the identified prioritised events. There are three likely situations: the proposed assessment might rank the identified situation as perfect as the ground truth, or it might rank it as not perfect as the ground truth, or, in the worst scenario, the system might rank different classes of activities as opposed to the ground truth.

Therefore, due to uncertainty, the modelling process will introduce an additional five different objects for the underlying scenario. The first object is the activity of interest and it is defined as a set in equation 5.10 as follows:

$$AoI_{\vartheta} \in \{aoi_1, aoi_2 \dots aoi_{\eta}\} \tag{5.10}$$

where the set $\vartheta \in AoI_{\beta}$, $\vartheta | 1 \leq \vartheta \leq \eta$

The second object is the normal activity and it is defined in equation 5.11 as follows:

$$A_{\varphi} = \left\{a_1, a_2 \dots a_{\zeta}\right\} \tag{5.11}$$

where the set $\varphi \in A_{\delta} \;\; \varphi | 1 \leq \varphi \leq \zeta$

If the SWA system is proposing different ranking paradigms for the identified events in comparison to the ground truth, the permutation concerning their degree of importance will also change accordingly. Thus, the underlying process will define the

degree of importance for different classes of tracking activities as a set; this is shown in equation 5.12 as follows:

$$DAoI_{\vartheta} = \{daoi_1, daoi_2 \ldots daoi_{\eta}\} \tag{5.12}$$

where the $\vartheta \in DAoI_{\varsigma} \quad \vartheta | 1 \leq \vartheta \leq \eta$

Moreover, due to capability issues with the SWA system, the proposed assessment component defines the ranking paradigms for the degree of importance concerning the normal activities as a set; this is defined in equation 5.13:

$$DA_{\varphi} = \left\{da_1, da_2 \ldots da_{\zeta}\right\} \tag{5.13}$$

Where the random of set $\varphi \in DA_{\Delta} \quad \varphi | 1 \leq \varphi \leq \zeta$

In regards to this, the *Proposed Assessment P.A* for the underlying situation is defined in equation 5.14 below as:

$$P.A_{\Gamma} = \{\{AoI_{\vartheta} \ni DAoI_{\vartheta}\} \cup \{A_{\varphi} \ni DA_{\varphi}\} \cup \{AE_{\alpha}\}\} \tag{5.14}$$

where $1 \leq \vartheta \leq \eta$, $1 \leq \varphi \leq \zeta$, $1 \leq \alpha \leq \sigma$, $1 < \Gamma < N \mid \Gamma = \eta + \zeta + \sigma$

It is equally important to mention that $AoI_{\beta} \ni DAoI_{\vartheta}$ and $A_{\delta} \ni DA_{\Delta}$. Hence, while the SWA system is experiencing ranking capabilities issues, the underlying system may not rank the AoI_{β} *and* A_{δ} as perfect as the ground truth. Therefore, if the ranking paradigms for different classes of events change, the permutation in relation to their degree of importance also changes accordingly.

This section has introduced the modelling scheme for situational assessment where the real time system is reporting different classes of tracking activities during a real time operation. Furthermore, we have defined three different components to represent

the SWA system output in the form of a list of tracking activities. The first component is the *Situational Assessment* objects, the second is *Ground Truth* objects and the third is *Proposed Assessment* objects.

The next section in the development phase is the *Scheduling Capability Score* (SCS) for evaluating the ranking capability of the real time system.

5.4 Developing Phase for the Scheduling Capability Score (SCS)

This section discusses the development phase of the " *Scheduling Capability Score*". The underlying performance metrics are intended to evaluate the ranking capability of the real time system in terms of shifting or scheduling important classes of tracking activity over the normal activities. Furthermore, the development phase is divided into three phases with the first phase computing the scheduling capability for the proposed assessment output during the real time operation. The evaluation process will compute the Current Capability Score CCS; this is defined in equation 5.15:

$$CCS = \sum_{\vartheta=1}^{\eta} \begin{pmatrix} (PA_{\Gamma} - 1 \mapsto AoI_{\vartheta}) + \eta_{MAX} \\ \eta_{MAX} \end{pmatrix} \tag{5.15}$$

where, $\vartheta \in AoI_{\beta} \mid 1 \leq \vartheta \leq \eta$, η= total number of important activity, $PA_{\Gamma} = \Gamma$, $1 \leq \Gamma \leq N \mid N$= The total number of tracking activity being proposed by real time system.

We have defined the CCS to compute the current scheduling state concerning the desired class of tracking activities, specifically it is measuring how the situational awareness system is shifting the activity of interest AoI among other identified events being proposed by the real time system.

However, during real time operation the proposed number of tracking activities is expected to go through one or more filtering processes such as the information perception, comprehension and projection, therefore the proposed size of different tracking activities it may be changed, these changes can directly impact the accuracy of the CCS scoring scheme. Therefore we have defined the *Actual Capability Score*; this is defined in equation 5.16 and the *Worst Capbility Score*; this is defined in equation 5.17 for each time the proposed number of tracking activity being filtered in a dynamically monitored environment.

$$ACS = \sum_{\beta=1}^{\eta} \begin{pmatrix} \left(GT_{\iota} - 1 \mapsto AoI_{\beta}\right) + \eta_{MAX} \\ \eta_{MAX} \end{pmatrix} \tag{5.16}$$

where

$1 \leq \beta \leq \eta$, $1 \leq \iota \leq \Psi$, $\Psi = \eta + \zeta$, $\mid \zeta =$ The total number of normal activity, and η=The total number of tracking activity in which the SWA system is regarded as an important event. And $GT_{\iota} = \iota$

$$WCS = \sum_{\beta=1}^{\eta} \begin{pmatrix} \left(NoH_{i} \mapsto AoI_{\beta}\right) + \eta_{MAX} \\ \eta_{MAX} \end{pmatrix} \tag{5.17}$$

where, $1 \leq \beta \leq \eta$, $1 \leq i \leq N \mid N =$ The total number of tracking activities being proposed by real time system. $\eta =$ The total number of tracking activities in which the SWA system is regarded as an important event.

The second phase introduces the *Scheduling Capability Score SCS*; this is defined in equation 5.18:

$$SCS = \frac{\sum_{\vartheta=1}^{\eta} \begin{pmatrix} (PA_{\Gamma} - 1 \mapsto AoI_{\vartheta}) + \eta_{MAX} \\ \eta_{MAX} \end{pmatrix}}{\sum_{\beta=1}^{\eta} \begin{pmatrix} \left(NoH_{i} \mapsto AoI_{\beta}\right) + \eta_{MAX} \\ \eta_{MAX} \end{pmatrix}} \tag{5.18}$$

where,

The random set of $\vartheta \in AoI_{\beta} \mid 1 \leq \vartheta \leq \eta$ and $1 \leq \beta \leq \eta \mid \eta =$ the total number of tracking activities in which the SWA system is regarded as an important event. $PA_{\Gamma} = \Gamma$, $1 \leq \Gamma \leq N \mid N=$ The total number of tracking activity being proposed by real time system. $NoH_i = N - i$ and $1 \leq i \leq N$

We have defined the SCS in equation 5.18 to quantify the scheduling process on information perception comprehension and projection regardless any changes to proposed number of tracking activity being proposed by the situational awareness system. It should provide a representative score for the user perception in comparison to the CCS defined in equation 5.15.

The third phase will normalise the SCS scoring scheme between [0-1]; this is to overcome the knowledge representation problem previously discussed in Chapter 4. The phase is divided into two stages. The first stage introduces the Good Scheduling State (GSS); this is defined in equation 5.19:

$$GSS = \frac{\sum_{\vartheta=1}^{\eta} \begin{pmatrix} \left(GT_{\iota} - 1 \mapsto AoI_{\beta}\right) + \eta_{MAX} \\ \eta_{MAX} \end{pmatrix}}{\sum_{\beta=1}^{\eta} \begin{pmatrix} \left(NoH_i \mapsto AoI_{\beta}\right) + \eta_{MAX} \\ \eta_{MAX} \end{pmatrix}} \tag{5.19}$$

where,

The random set of $\vartheta \in AoI_{\beta} \mid 1 \leq \vartheta \leq \eta$ and $1 \leq \beta \leq \eta \mid \eta =$ The total number of tracking activities in which the SWA system is regarded as an important event. $1 \leq \iota \leq \Psi$, $\Psi = \eta + \zeta$, $\mid \zeta=$ The total number of normal activities, $GT_{\iota} = \iota$, $NoH_i = N - i$ and $1 \leq i \leq N \mid N=$ The total number of tracking activities being proposed by the real time system.

The second stage introduces the *Scheduling Capability Score′* SCS' defined in equation 5.20:

$$SCS' = \begin{cases} if\ SCS \neq GSS \\ \left(\dfrac{\sum_{\vartheta=1}^{\eta} \binom{(PA_{\Gamma} - 1 \mapsto AoI_{\vartheta}) + \eta_{MAX}}{\eta_{MAX}}}{\sum_{\beta=1}^{\eta} \binom{\left(NoH_i \mapsto AoI_{\beta}\right) + \eta_{MAX}}{\eta_{MAX}}} \right) \\ \\ if\ SCS = GSS \\ \left(\dfrac{\sum_{\vartheta=1}^{\eta} \binom{\left(GT_{\iota} - 1 \mapsto AoI_{\beta}\right) + \eta_{MAX}}{\eta_{MAX}}}{\sum_{\beta=1}^{\eta} \binom{\left(NoH_i \mapsto AoI_{\beta}\right) + \eta_{MAX}}{\eta_{MAX}}} \right) - \left(\dfrac{\sum_{\vartheta=1}^{\eta} \binom{(PA_{\Gamma} - 1 \mapsto AoI_{\vartheta}) + \eta_{MAX}}{\eta_{MAX}}}{\sum_{\beta=1}^{\eta} \binom{\left(NoH_i \mapsto AoI_{\beta}\right) + \eta_{MAX}}{\eta_{MAX}}} \right) \end{cases} \tag{5.20}$$

The next section examines the proposed performance metrics against its intended purpose with two levels of assessments.

5.5 Evaluation Process

This section examines the proposed performance metric against its intended purpose with two levels of assessments, to challenge the *Scheduling Capability Score′* SCS'. The first level is a case study based evaluation used to guide researchers from various disciplines in how to adopt proposed metrics to their domain specific needs and configuration.

The second level is a quality-based evaluation to examine the underlying metric against its intended purpose. This level has been designed to verify the scalability of the proposed performance metric over three separate scenarios. Each scenario encompasses different numbers of tracking activities which are regarded as important events by the real time system.

5.5.1 Case Study Based Evaluation

This section demonstrates an example used as guidance for researchers from various disciplines in adopting the proposed performance metric to their domain specific configuration. This section is divided as follows: the first part substitutes the situational assessment objects which have been defined in equation 5.8; the second part quantifies the scheduling capability of the real time system using the *Scheduling Capability Score′ SCS′* defined in section 5.2.

Substitutes the Situational Assessment Objects

According to the case study in section(5.2), the real time system has identified different classes of tracking activity, as shown in Table 5.3. Due to capability issues with the lower level of the situational awareness system, the perception and comprehension stage has not ranked or scheduled different classes of tracking activity in accordance to their degree of importance.

Moreover, to evaluate the scheduling capability for the underlying situational assessment, we have defined 7 different objects in section 5.3. This is to model the ranking capability problem of a real time system and allows researchers from various disciplines to use the proposed modelling scheme for their domain specific configuration. Thus, this section substitutes the situational assessment objects concerning the case study demonstrated in section 5.2.

The first object is the tracking activity in which the real time system is regarded as an important event AoI_{β}; this is defined in equation 5.1:

$AoI_{\beta} = aoi_1(1), aoi_2(2)$, where $1 \leq \beta \leq \zeta |\ \eta{=}2$

The second object is the tracking activity in which the real time system is regarded as normal activity A_{δ}; this is defined in equation 5.2:

$A_{\delta} = a_1(3), a_2(4), a_3(5), a_4(6)$, where $1 \leq \delta \leq \zeta \mid \zeta{=}2$

The third object is the undesired events AE_{α} and is defined in equation 5.3:

AE_{α}= $ae_1(1), ae_2(2)$, where $1 \leq \alpha \leq \sigma \mid \sigma$=2

The fourth object is the priority score for the activity of interest $DAoI_{\varsigma}$; this is defined in equation 5.4:

$DAoI_{\varsigma} = daoi_1(2), daoi_2(1)$

The fifth object is the score of importance SoI_j for all tracking activity proposed by the real time system; this is defined in equation 5.5:

$SoI_j = soi_1(8), soi_1(7), soi_1(6), soi_1(5), soi_1(4), soi_1(3), soi_1(2), soi_1(1)$, where, $1 \leq j \leq N \mid N = 8$.

The sixth object is the predetermined order NoH_h for all the tracking activity being proposed by real time system and is defined in equation 5.6:

$NoH_i = noh_1(7), noh_2(6), noh_3(5), noh_4(4), noh_5(3), noh_6(2), noh_7(1), noh_8(0)$, where, $1 \leq h \leq N \mid ToH = 8 - 1 = 7$.

The final object is the priority score for the normal activity DA_{Δ} and is defined in equation 5.7:

$DA_{\Delta} = da_1(4), da_2(3), da_3(2), da_4(1)$

The next section quantifies the scheduling capability for the multilevel situational assessment using the proposed performance metric.

Measuring the Scheduling Capability of Real Time System

This section demonstrates the evaluation methodology to quantify the scheduling capability of the real time system. The evaluation process will be divided into two phases with the first phase measuring the scheduling capability during the perception stage, as shown in Table 5.1. The second phase computes the scheduling capability during the comprehension stage as shown in Table 5.2.

The next section evaluates the ranking capability during the perception stage using the *Scheduling Capability Score SCS*.

Quantifying the Ranking Capability on Information Perception and Comprehension

In order to measure the capability of the situational assessment at the perception stage, we start by computing the current capability score as shown in equation 5.15:

$$CCS = \sum_{\vartheta=1}^{\eta=2} \begin{pmatrix} (PA_\Gamma - 1 \mapsto AoI_1) = 2 + (\eta_{MAX}) = 4 \\ (\eta_{MAX}) = 4 \end{pmatrix} + \begin{pmatrix} (PA_\Gamma - 1 \mapsto AoI_2) = 4 + (\eta_{MAX}) = 4 \\ (\eta_{MAX}) = 4 \end{pmatrix}$$
$$= \binom{6}{4} + \binom{8}{4} = 85$$

The second step computes the actual capability score, as shown in equation 5.16 ACS:

$$ACS = \sum_{\beta=1}^{\eta=2} \begin{pmatrix} (GT_\iota - 1 \mapsto AoI_1) = 0 + (\eta_{MAX}) = 4 \\ (\eta_{MAX}) = 4 \end{pmatrix} + \begin{pmatrix} (GT_\iota - 1 \mapsto AoI_1) = 1 + (\eta_{MAX}) = 4 \\ (\eta_{MAX}) = 4 \end{pmatrix} =$$
$$\binom{4}{4} + \binom{5}{4} = 6$$

The third step computes the worst capability score $WCS =$, as shown in equation 5.17:

$$WCS = \sum_{\beta=1}^{\eta=2} \begin{pmatrix} (NoH_1 \mapsto AoI_1) = 7 + (\eta_{MAX}) = 4 \\ (\eta_{MAX}) = 4 \end{pmatrix} + \begin{pmatrix} (NoH_2 \mapsto AoI_2) = 6 + (\eta_{MAX}) = 4 \\ (\eta_{MAX}) = 4 \end{pmatrix}$$
$$= \binom{11}{4} + \binom{10}{4} = 540.$$

The fourth step computes the scheduling capability score SCS as shown in equation 5.18:

$$SCS = \frac{\sum_{\vartheta=1}^{\eta} \begin{pmatrix} (PA_\Gamma - 1 \mapsto AoI_\vartheta) + \eta_{MAX} \\ \eta_{MAX} \end{pmatrix}}{\sum_{\beta=1}^{\eta} \begin{pmatrix} \left(NoH_h \mapsto AoI_\beta\right) + \eta_{MAX} \\ \eta_{MAX} \end{pmatrix}} = \frac{\binom{6}{4} + \binom{8}{4}}{\binom{11}{4} + \binom{10}{4}} = \frac{85}{540}$$

The fifth step computes the good scheduling state as shown in equation 5.19:

$$GSS = \frac{\sum_{\vartheta=1}^{\eta} \binom{\left(GT_{\iota} - 1 \mapsto AoI_{\beta}\right)}{\eta_{MAX}}}{\sum_{\beta=1}^{\eta} \binom{\left(NoH_{h} \mapsto AoI_{\beta}\right) + \eta_{MAX}}{\eta_{MAX}}} = \frac{\binom{4}{4} + \binom{5}{4}}{\binom{11}{4} + \binom{10}{4}} = \frac{6}{540}$$

Finally, we compute the *Scheduling Capability Score'* SCS' during the perception stage as follows:

$$SCS' = \frac{\binom{6}{4} + \binom{8}{4}}{\binom{11}{4} + \binom{10}{4}} = \frac{85}{540} = 0.157$$

Then, during the comprehension stage as follows:

$$SCS' = \frac{\binom{5}{4} + \binom{6}{4}}{\binom{11}{4} + \binom{10}{4}} = \frac{20}{540} = 0.037$$

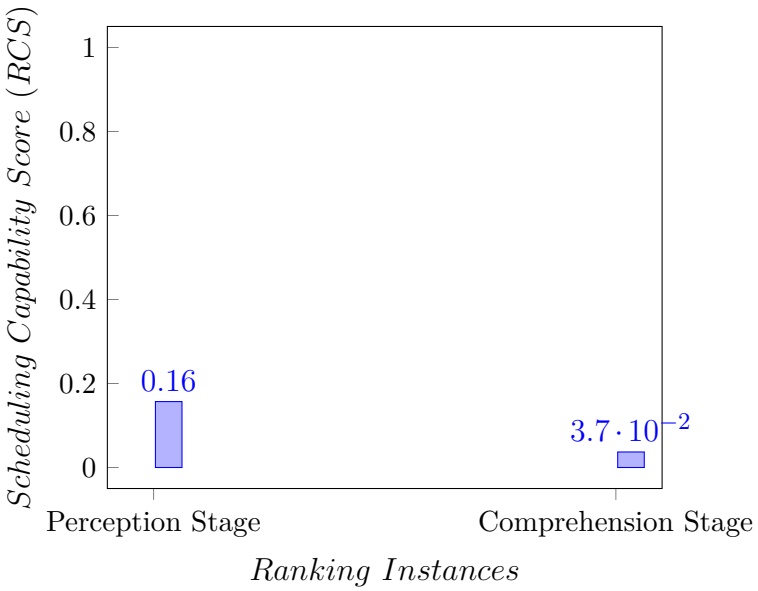


Fig. 5.1 Case Study Based Evaluation: Quantifying the Ranking Capability of Real Time System using the Scheduling Capability Score *SCS*

Based on the above scenario, the proposed assessment at the comprehension stage scores better than at the perception stage; the SWA awareness at the perception stage scores (0.157) meaning the perception stage has not ranked the important class of tracking activity as perfect as the ground truth. The proposed assessment at the comprehension stage scores (0.037). Similarly, the SWA has ranked the activity of

interest as not perfect as the ground truth. However, the ranking paradigm at the comprehension stage is better than at the perception stage. Resultantly, the underlying metrics have quantified the scheduling capability for a multilevel situational assessment appropriately.

Furthermore, the case study based evaluation does not examine the capability of the proposed performance metric, in terms of providing an evidencing score for all the ranking instances concerning the scheduling process of the real time system. Rather, it has successfully quantified only two ranking instances for the underlying scenario. The next section conducts a quality based evaluation to examine the Scheduling Capability Score SCS against all ranking instances over three separates scenarios.

5.5.2 Quality Based Evaluation

This section examines the proposed metric against its intended purpose. The Scheduling Capability Score is expected to provide unique scoring scheme for all the ranking instances concerning with scheduling process of real time system. This is from the good state to the worst state, respectively.

The evaluation process will validate the proposed metrics against its intended purpose over three separate scenarios. In the first scenario the real time system has regarded two tracking activities as important from all the events being proposed; this has been demonstrated in section 5.2. The second scenario regards three tracking activities as important and, finally, the third scenario regards four tracking activities as important. It is important to mention that we have designed these scenarios based on the two inputs for the combination operation n choose k. We have assumed the size of n=8 for simplicity reason, and following the case study demonstrated in existing literatures [74],[77]. Furthermore we have assumed different sizes for the k term based

on the determination point we have discussed in chapter 3, section 3.5.2 Determination Point for the Maximum Number of AoIs.

Furthermore, validation is required to simulate all the ranking instance concerning the scheduling process over the three separates scenarios.

The real time system tracked 8 different activities during the first scenario with two of them regarded as important. Additionally, the scheduling capability score is expected to provide a unique scoring scheme for all the ranking instances. With this in mind, we computed the number of ranking instances using the combination equation before running a numerical simulation using Matlab to mimic all the ranking instances. Next, we configured the Scheduling Capability Score SCS to measure all the ranking instances for the first scenario; the obtained results are shown in Figures 5.1 and 5.2.

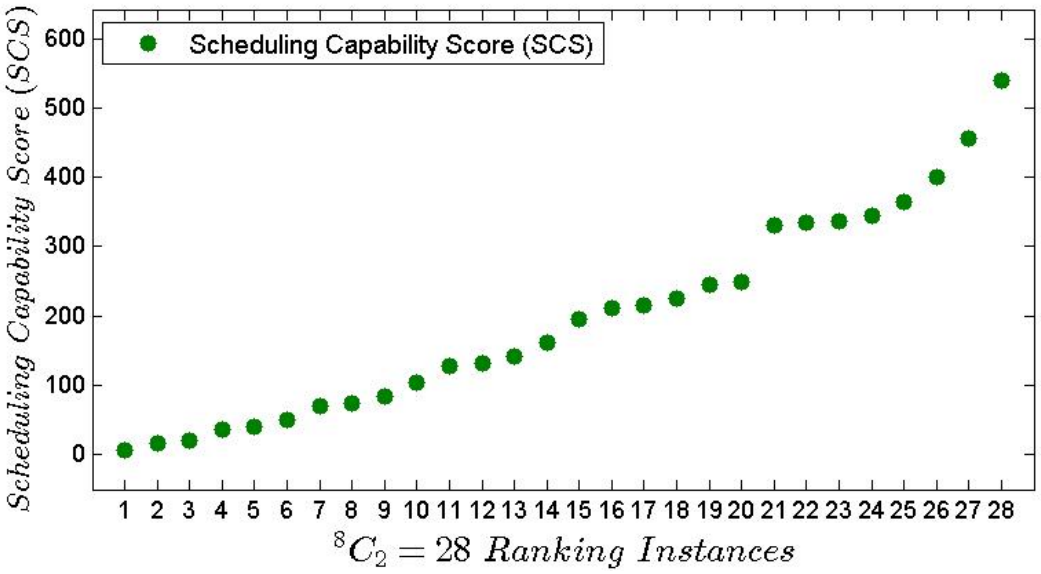


Fig. 5.2 Scenario 1 : Quality Based Evaluation For Validating The *Scheduling Capability Score* In Term of Providing Unique Scoring Scheme for All The Ranking Instances Obtained By The Combination Operation $^8C_2 = \frac{8!}{2!(8-2)!}$.

During the second scenario, we increased the number of important events up to three tracking activities; the obtained results are shown in Figures 5.3 and 5.4.

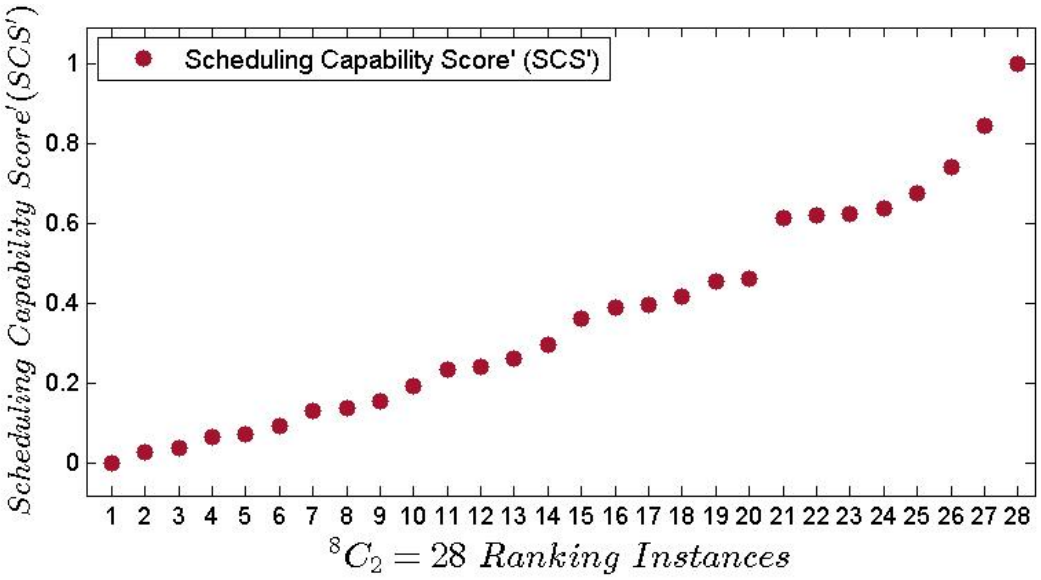


Fig. 5.3 Scenario 1 :Quality Based Evaluation For Validating The *Scheduling Capability Score′* In Term of Providing Unique Scoring Scheme For All The Ranking Instances Obtained By The Combination Operation $^8C_2 = \frac{8!}{2!(8-2)!}$.

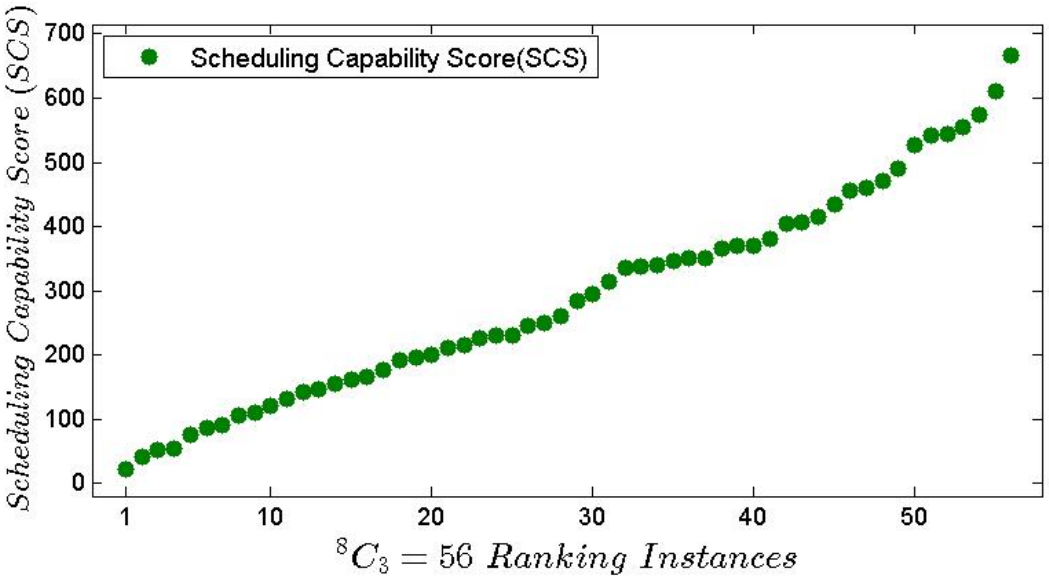


Fig. 5.4 Scenario 2 : Quality Based Evaluation For Validating The *Scheduling Capability Score* In Term of Providing Unique Scoring Scheme For All The Ranking Instances Obtained By The Combination Operation $^8C_3 = \frac{8!}{3!(8-3)!}$.

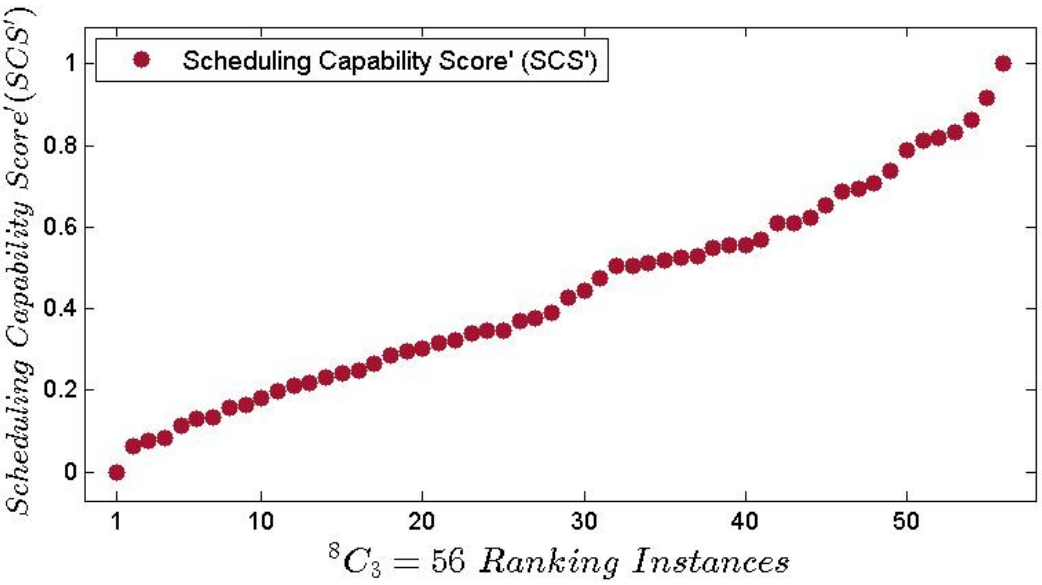


Fig. 5.5 Scenario 2 : Quality Based Evaluation For Validating The *Scheduling Capability Score′* In Term of Providing Unique Scoring Scheme For All The Ranking Instances Obtained By The Combination Operation $^8C_3 = \frac{8!}{3!(8-3)!}$.

Finally, during the third scenario we increased the number of important activities to our determination point, where the real time system is expected to deal with a maximum of four tracking activities during the real time operation. The obtained results are shown in Figures 5.5 and 5.6.

This section examined the Scheduling Capability Score against its intended purpose over three separate scenarios and it was found the proposed performance metric successfully quantified the scheduling capability of the real time system. The proposed scoring scheme can notify the decision making resource about any ranking capability issue that may occur during a real time operation. The next section will provide a comparative evaluation using existing performance metrics.

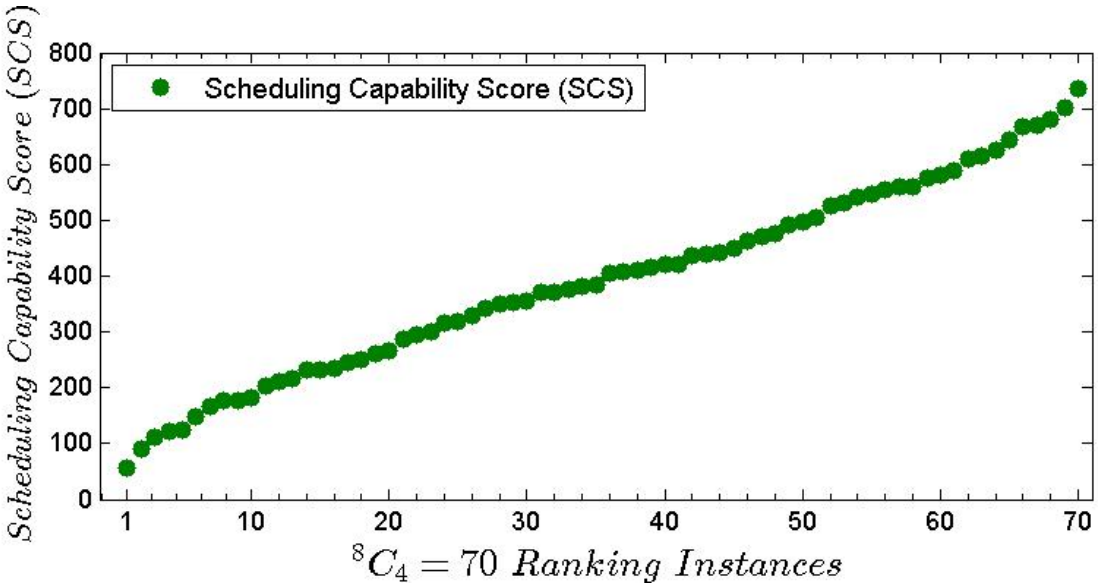


Fig. 5.6 Scenario 3 : Quality Based Evaluation For Validating The *Scheduling Capability Score* In Term of Providing Unique Scoring Scheme For All The Ranking Instances Obtained By The Combination Operation $^8C_4 = \frac{8!}{4!(8-4)!}$

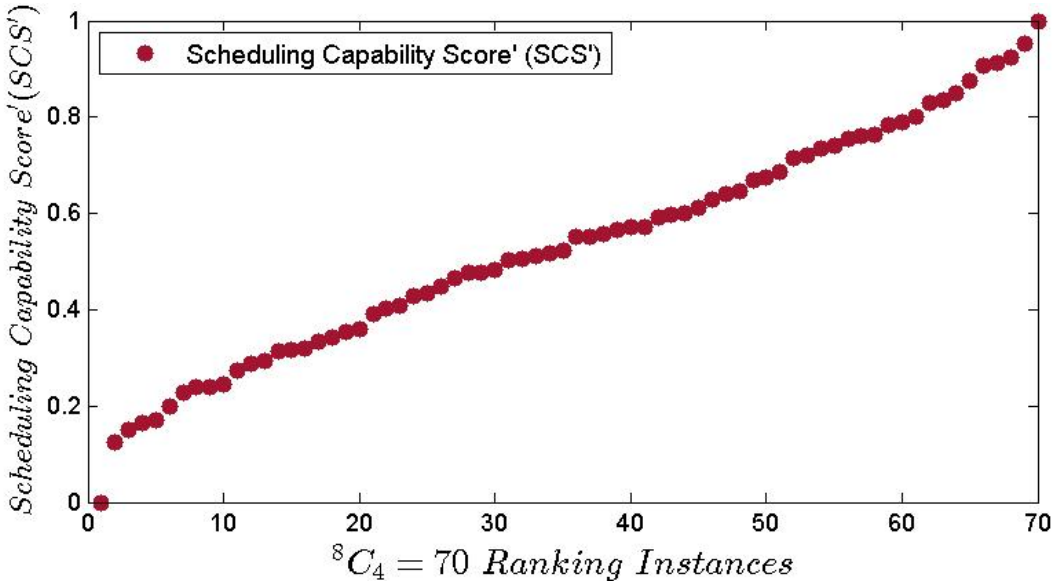


Fig. 5.7 Scenario 3 : Quality Based Evaluation For Validating The *Scheduling Capability Score′* In Term of Providing Unique Scoring Scheme For All The Ranking Instances Obtained By The Combination Operation $^8C_4 = \frac{8!}{4!(8-4)!}$.

5.6 Comparative Evaluation

This section conducts a quality based evaluation to examine the *Activity of Interest Score* in terms of providing an appropriate scoring scheme for all the ranking instances concerning the scheduling process. The obtained result is shown in Figure 5.8

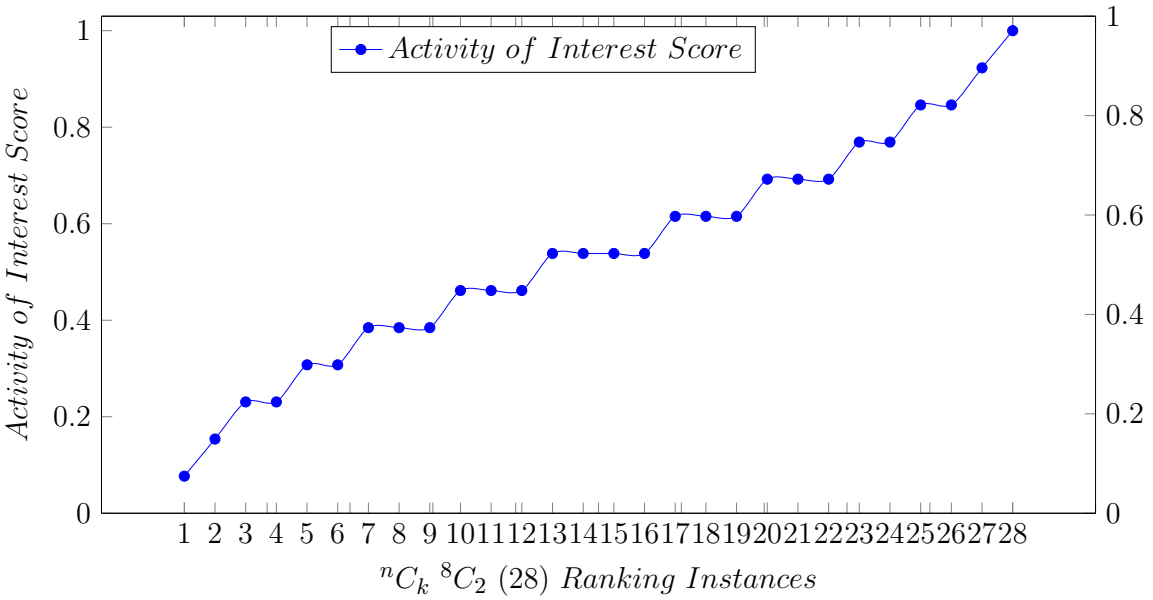


Fig. 5.8 Scenario 1 : Quality Based Evaluation for Validating *Activity of Interest Score*

Unfortunately, the underlying performance metric did not successfully quantify all the ranking instances concerning the scheduling process of the real time system. This means it had not been designed for measuring the scheduling process of the SWA domain. The next section discusses the computational complexity involved when evaluating the ranking capability of the real time system.

5.7 Computational Complexity

The number of computational steps can sometimes become an issue during a real time operation; this is caused by, but not limited to, the following problems: the

limitation of hardware resources being used for assessing emerging situations; the frequent occasions needed for quantifying different ranking instances in dynamically monitored environments; the number of computational steps required to complete the evaluation process.

The first two problems are either subjected to a domain specific scenario or to contextual situational assessments, while the last computational problem is directly related to the assessment methods for evaluating the ranking capability of the real time system. It is important to mention that this work is not interested in the computational complexity levels involved in response to the growth of inputs. Rather, it is intended to analyse the computational process concerning the two distinct operations utilised for the process refinement stage of the Joint Director of Laboratories (JDL); the first and second stages are assessment and optimisation, reespectively.

This section discusses the computational complexity involved in evaluating either the prioritisation process or scheduling tasks of the real time system.

5.7.1 Prioritisation Process

To understand the computational complexity concerning the number of steps, let us assume that the situational awareness, SWA, system is reporting a priority list with a number of distinct events, each of them with a different degree of importance. The notion of different ranking instances concerning the number of prioritised events is related to the act of rearranging, or permuting, all the identified events into some sequence or order. Furthermore, the number of permutations for each priority list can be determined by n factorial (usually written as $n!$); this is the product of all positive integers less than or equal to n. This allows us to define the number of ranking instances for each priority list using the operation of factorial (N!), where N represents the number of identified events, and factorial ! provides the number of all possible

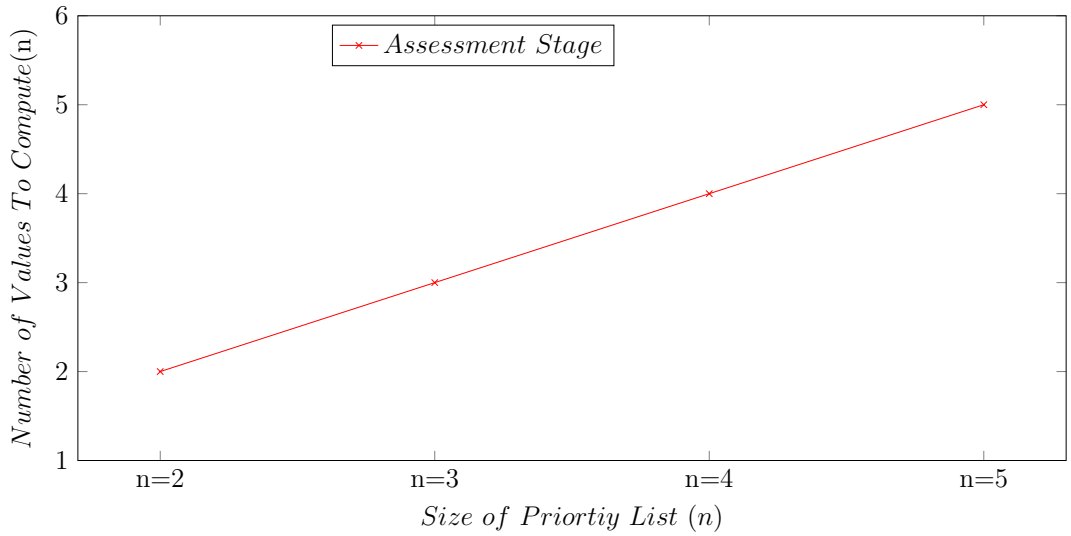


Fig. 5.9 The number of values required to compute for completion of the assessment stage, concerning the prioritisation process of the real time system

ranking instances for those activities. If the system has proposed a priority list with

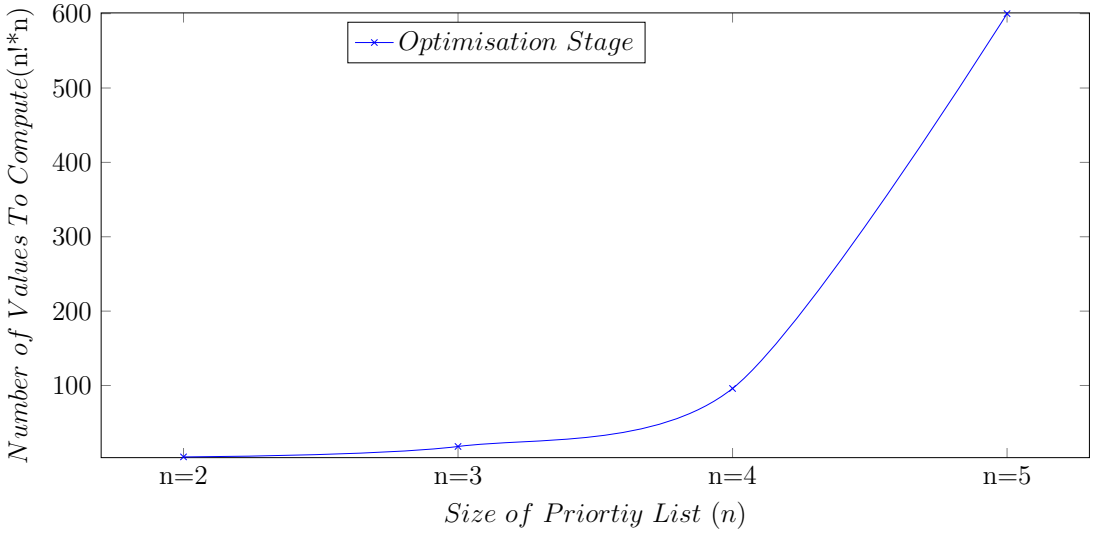


Fig. 5.10 The number of values required to compute for completion of the optimisation process concerning the prioritisation process of the real time system.

two complex events, the total number of state is obtained by $(2!) = 2$ then we will have two ranking instances. Likewise, if there are three distinct events, then the number of ranking instances can be obtained by factorial $(3!) = 6$ ranking instances. It is

apparent that when the number of prioritised events for each priority list increased, the number of ranking instances also increased accordingly.

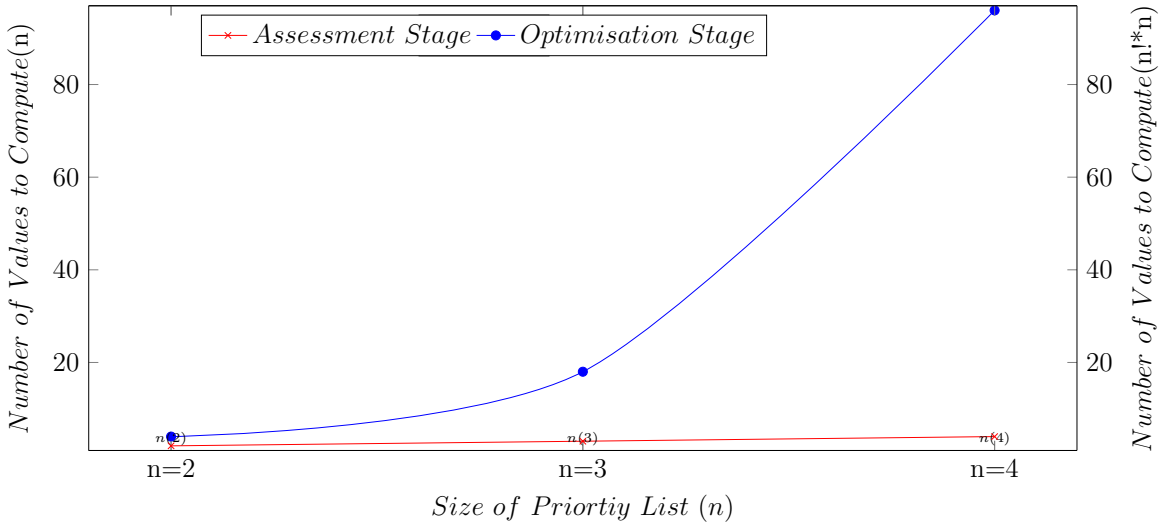


Fig. 5.11 Comparative Evaluation; Number of values to compute concerning the prioritisation process

The computational complexity involved in evaluating the ranking capability depends on two underpinning concepts of the process refinement stage. The first concept is the assessments stage, where the underlying method of evaluation is required to compute relevant values only for the proposed ranking instances; this is to verify the capability of the real time system. The second concept is the optimisation stage, where the underlying method of evaluation is required to compute all the ranking instances in relation to the proposed priority list. Hence, we assumed the worst scenario, where the computation process computes all the values for any given scenario; this is to enable the potential optimisation techniques to select the desired ranking paradigms for improving the ranking capability of the real time system.

Intuitively, during the assessment stage, the employed performance metric is required to compute a lesser number of values in comparison to the optimisation process. Likewise, the evaluation process is required to compute only the necessary values for assessing the prioritisation process of the real time system. However, during the

optimisation operation, the potential performance metric is required (at least) to compute the values for all ranking instances being obtained by N! operation.

For example, if the system is proposing a priority list with n number of events, the assessment stage requires only to compute n number of values in order to quantify the ranking capability. However, in the optimisation stage the underlying process is required to compute the $(n!) \times n$ number values to complete the evaluation process for any given scenario.

Initially, it seems the computation complicity between the two operations is very noticeable: this is because the factorial operation in the O big notation O(N!) is much more complex than the $O(N)$ terms. In fact, because we are dealing only with high abstract views of multidisciplinary areas, the number of prioritised events for each priority list is relatively very small, in comparison to lower levels. Furthermore, in previous works [79], [80] we have defined a determination point for the growth of N inputs. This allowed the evaluation process to be limited to an input size of four prioritised events for each priority list, concerning three dimensional views for what is deemed an important event. We have discussed the growth of n input with more details in chapter 3, section 3.5.2 *Determination Point for the Maximum Number of AoIs*.

With this in mind, there are three likely scenarios that might occur during real time operation. The situational awareness system might propose a priority list with various length of 2, 3 *or*, 4 complex events; these are the potential inputs for the evaluation operation. Therefore, having the knowledge of what to expect, in respect to the sizes and limits for each priority list, we can easily analyse the computational complexity problem without the fear of n term growth for the evaluation method.

The task of analysis become easier now, since we have already defined the O big notation for the optimisation stage O$((n!) \times n)$ and assessments process $O(N)$. It is

important to mention that this work is not interested in the computational complexity levels involved in response to the growth of inputs; rather it is intended to analyse the computational process during two distinct operations: the assessment operation and optimisation process. We have defined a method to analyse the computational complexity in terms of number of values required for assessing the prioritisation process as follows:

$$\mathrm{NAPP} = \mathrm{N} \tag{5.21}$$

where

$NAPP$ = Number of values required to be computed during the assessment stage, concerning with prioritisation process

N = Number of events being proposed by real time system

Furthermore, we have defined a method to analyse the computational complexity involved during the optimisation stage as follows:

$$\mathrm{NOPP} = \mathrm{N}! \times \mathrm{N} \tag{5.22}$$

where

$NOPP$ = Number of values required to be computed during the optimisation stage, concerning with prioritisation process

$N!$ = Number of all the ranking instances concerning the prioritisation process

N = Number of events being proposed by real time system

Equations 5.21 and 5.22 are subjected to two different operations concerning the prioritisation process. The first task is the assessment stage: here the evaluation

method is required to compute only the n number of values. For example, consider a scenario where the real-time system is reporting a priority list of two complex events; each of them has a different degree of importance.

Initially, the evaluation method is required to compute only two values for assessing the ranking paradigms of each complex event. With a priority list of three complex events, the performance metric will compute at least three values for assessing the situation. Therefore, the number of values for computing will increase with the increased length of each priority list. Similarly, for the assessment stage, the evaluation methods will compute the number of values related only to the perceived ranking instances and this is shown in figure 5.9. However, in the optimisation stage, the performance metric is required to compute n values for all the ranking instances which it can obtain by $(N!)$ operation and this is shown in figure 5.10.

With this in mind, we can substitute the first constant for the assessment stage with n number of values and the size of n depends on the length of each priority list, while in the optimisation stage we can substitute it with $n \times N!$. This computes the necessary values for all ranking instances concerning the perceived priority list. A method has been developed to compare the computational complexity, to complete the two different operations of evaluations. Therefore, we can use equation 5.21 for computing the computational complexity for the assessment stage, and equation 5.22 for quantifying the computational complexity of the optimisation process.

The next section discusses the computational complexity involved in evaluating the ranking capability concerning the scheduling process of the real time system.

5.7.2 Scheduling Process

To understand the computational complexity concerning the scheduling process, let us assume that the situational awareness, SWA, system is reporting a priority list with

two different classes of events, each of them with a different degree of importance. The notion of different ranking instances concerning the scheduling of relevant groups or classes of events over another undesired one is related to the act of combinatorial. This is to select or shift a k number of events over another undesired one in a list of n prioritised events, which has been proposed by a real-time system. Therefore,

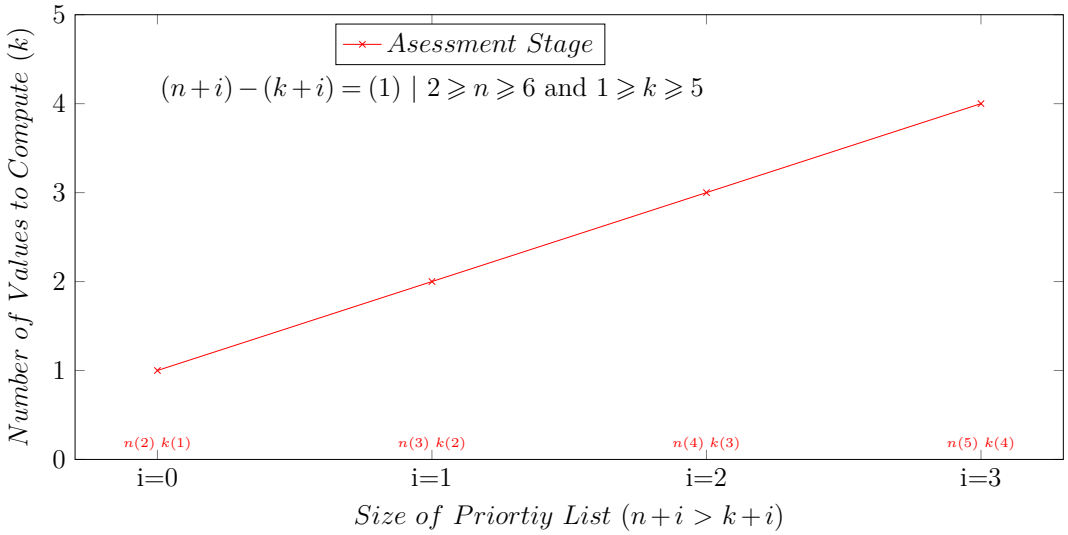


Fig. 5.12 The number of values required to compute in order to complete the assessment stage, concerning the scheduling process of real time system

the number of combinations for all elements concerning the desired class of event k in a given list of priority n, can be determined by n choose k (usually written as ${}^{n}C_{k}$), which means the combination of n things taken k at a time without repetition. That allows us to define the number of ranking instances for each priority list using the operation of combination (${}^{n}C_{k}$), where N represents the number of identified events for the proposed priority lists, and K represents the number of desired class of events. If the system has proposed a priority list with $n = 8$ number of events, this encompasses two different classes of K_1, and k_2. The total number of states is obtained by the combination operation $({}^{8}C_{k1=3}) = 56$, which results in fifty-six different ranking instances concerning the scheduling process, likewise, for the second class of events

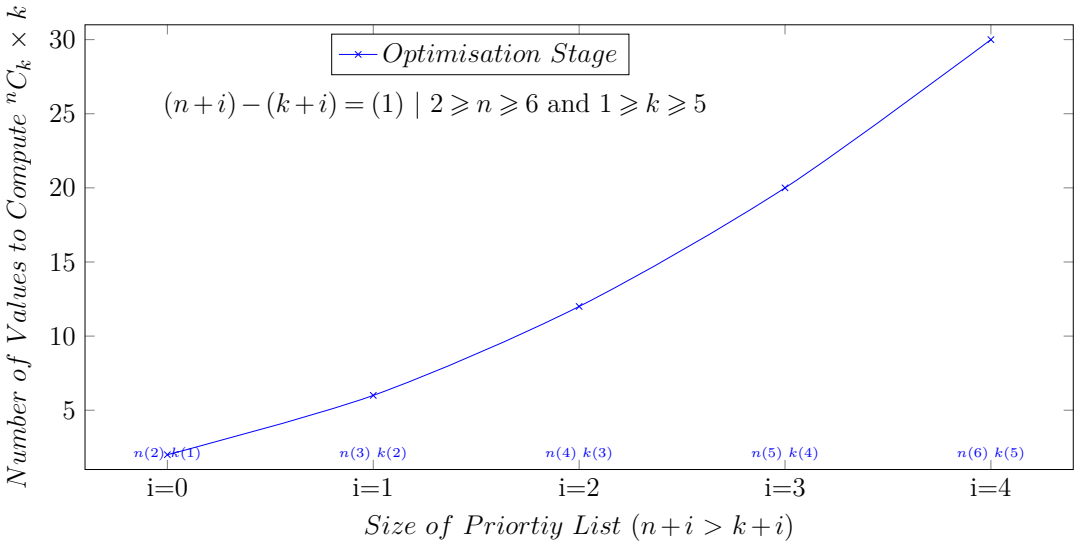


Fig. 5.13 The number of values required to compute for the completion of the optimisation stage, concerning the scheduling process of real time system

$k_2 = 5$, then the number of ranking instances can be obtained by the combination operation $({}^{8}C_{k2=5}) = 70$ ranking instances. It is essential to mention that the number of ranking instances is either decreased or increased in proportion to the changes of N or K sizes.

In other words, when the system is reporting a priority list with a large number of prioritised events, the number of ranking instances will increase. Consequently, in such a case, the number of values required to be computed will also increase. Moreover, the possible combination of ranking instances are also increased proportional to the difference in numbers between n and K sizes. Likewise, when the difference in numbers between n and k decreased, the number of ranking instance decreased proportionally and vice versa.

Therefore, during the assessment stage, the evaluation method is required to compute k number of values to verify the scheduling capability of the real time system. In such a case, the size of k term can influence the computational complexity for evaluating the scheduling process of the real time system, if and only if $2 \geqslant n \geqslant S$

where $S \geqslant 2$ and $K \geqslant 1 \mid n > K$. With this in mind, we have defined a method to analyse the computational complexity required for assessing the scheduling process as follows;

$$\text{NASP} = \text{K} \tag{5.23}$$

where

$NASP =$ Number of values required to be computed during the assessment stage, concerning the scheduling process

$K =$ Number of events concerning the desired classes of events

Similarly, during the optimisation process, we have assumed the worst case scenario, where the evaluation methods are required to compute a k number of values for all the ranking instance ${}^{N}C_{K}$, that is to enable the potential optimisation techniques to select the desired state among all others. In such cases, the optimisation operation is influenced by two terms; the number of all the events being proposed by the real time system, in which is defined equation 5.24 as N and the number of events concerning the desired classes of events K. Therefore, we have defined a method to analyse the computational complexity involved during the optimisation stage as follows:

$$\text{NOSP} = {}^{N}C_{K} \times K \tag{5.24}$$

where

$NOSP =$ Number of values required to be computed during the optimisation stage, concerning the scheduling process

$N =$ Number of all events being proposed by real time system

$K =$ Number of events concerning the desired classes of events

$^{N}C_{K}$ = Number of all the ranking instances concerning the scheduling process

Moreover, the computational complexity involved in the optimisation process is also influenced by the number of ranking instances concerned with the perceived priority list. Analytically, there are two different factors where the scheduling process can influence the number of ranking instances. The first is the k size for the desired class of events and the second is the difference in number between N and K sizes.

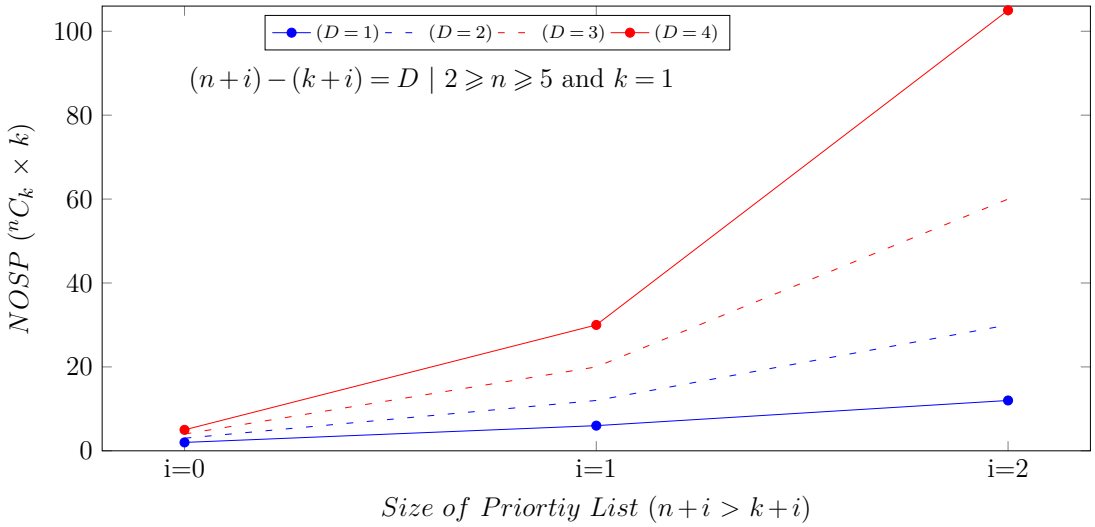


Fig. 5.14 Computational complexity for various sizes of priority lists, concerning the scheduling process

In Figure 5.12 when $i = 0$ the $^{n}C_{k}$ becomes $^{n=2}C_{K=1}$, where the optimisation stage is required to compute only $^{2}C_{1} \times 1 = (2)$ values for completing the evaluation operation. Furthermore, when $i = 1$ and the $^{n}C_{k}$ become $^{n=3}C_{K=2}$, the evaluation method is required to compute (6) values for assessing the scheduling process during the optimization stage. Similarly, when i=3, the number of values to be computed is (12).

Noticeably, when the i values increase, the k and n sizes also increase. Consequently, the computational complexity in terms of the number of values to be computed also

increases respectively. In such cases, we have observed the computational complexity of assessing the scheduling process during the optimisation stage, where the difference in number between n and k equals to (1). The evaluation process is intended to analyse how the d factor can influence the computational complexity concerning the scheduling process.

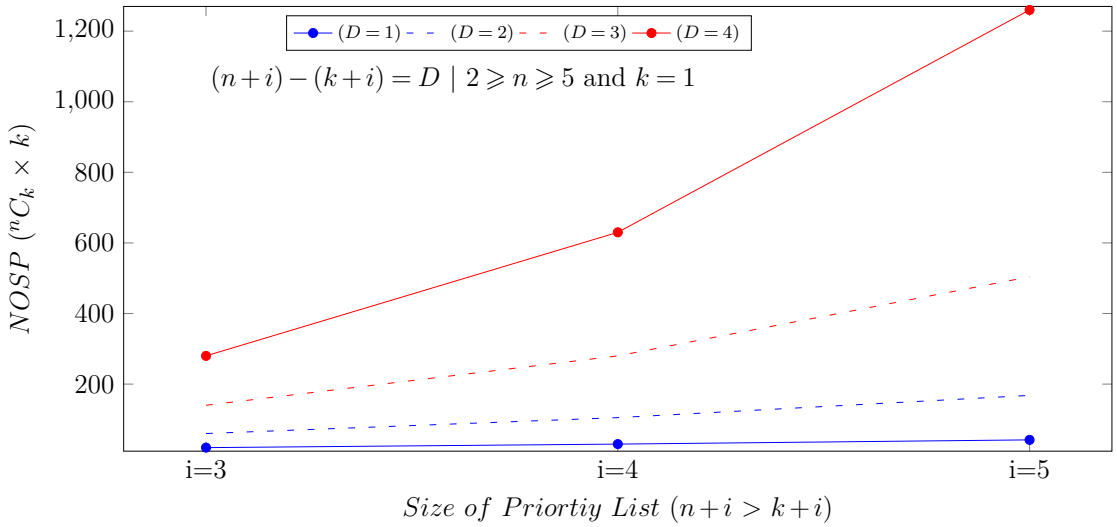


Fig. 5.15 Computational complexity for various sizes of priority lists, concerning the scheduling process

Thus far, we have decided to observe further the computational complexity under various sizes of the different priority list as shown in Figures 5.14 and 5.15 where the difference in numbers between n and k are changing in increasing order. Specifically, we have defined the following condition$(n+i)-(k+i) = D \mid 2 \geqslant n \geqslant 5$, and $k = 1$. That is, to allow the evaluation process to systemically observe the computational complexity over the increasing number of n and k over D.

Looking at Figure 5.14, when i=0 and D=1, the evaluation method is required to compute a lesser number of values in comparison to D=2, D=3, or D=4. Similarly, Figure 5.15 shows that regardless of size D, the increases in some prioritised events n and k can still influence the computational process during the optimisation stage, as

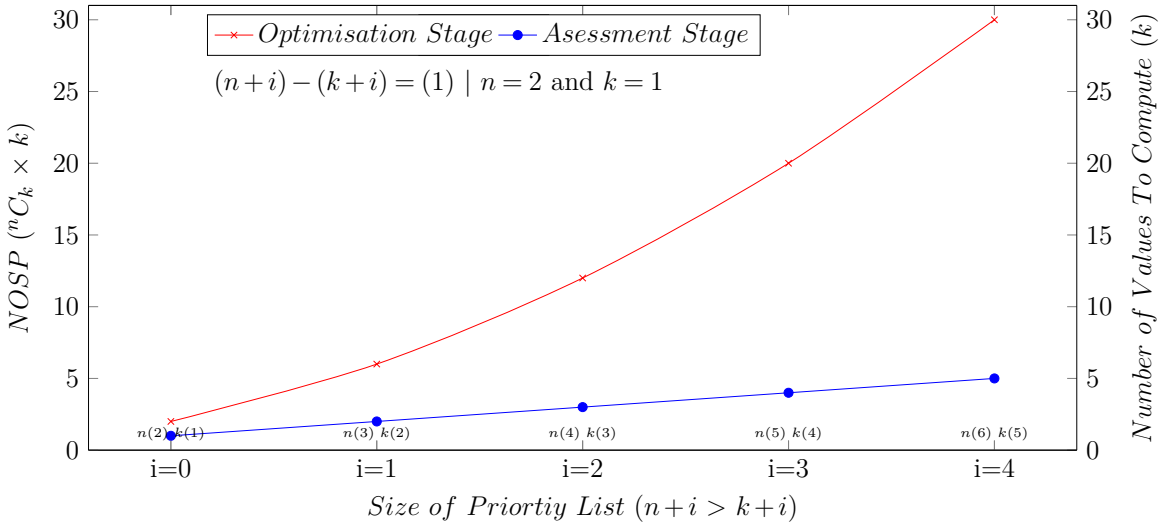


Fig. 5.16 Comparative Evaluation; Number of values to compute concerning the scheduling process, $(n > k)$=1

shown in Figure 5.16. However, the increased number in d term has greater influence than the I term.

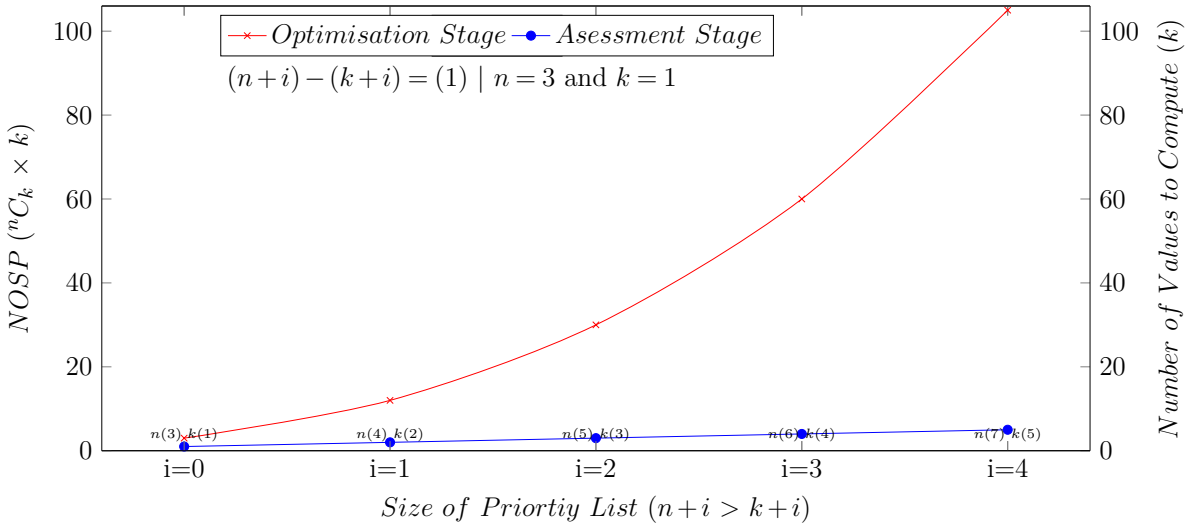


Fig. 5.17 Comparative Evaluation; Number of values to compute concerning the scheduling process, $(n > k)$=2

We conducted further analysis of the evaluation process to observe the computational complexity during the optimization stage in comparison to the assessment stage, concerning the increased input of terms I and D as shown in Figures 5.17 5.18 and

5.19. Furthermore, when $d = 1$ and $i = 4$, as shown in Figure 5.20, the evaluation method was required to compute at least (30) values for assessing the scheduling process during the optimisation stage and (5) values during the assessment stage. Similarly, during the time when the $d = 2$ and the $i = 4$, as shown in Figure 5.17,

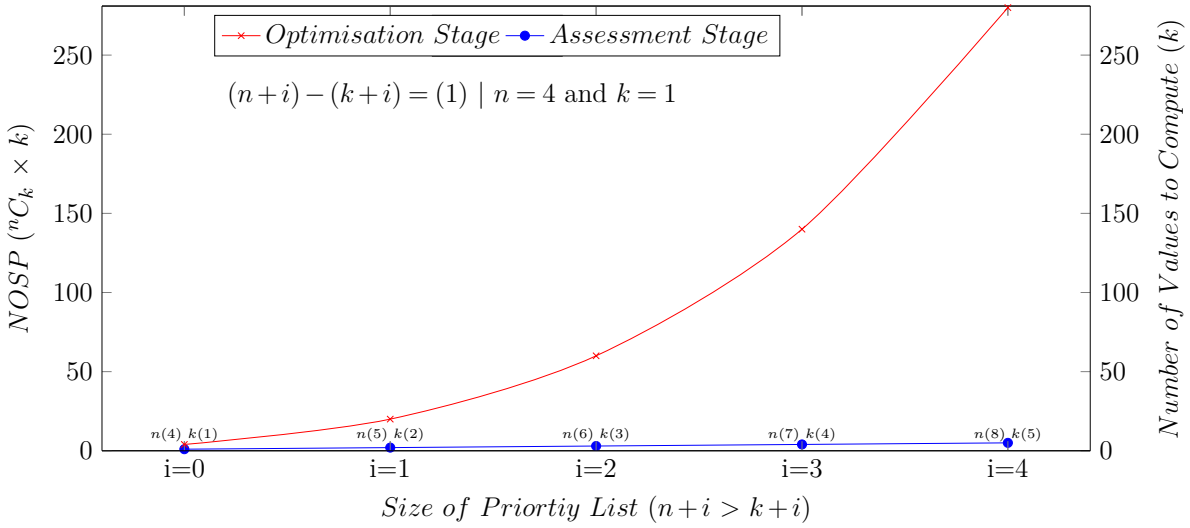


Fig. 5.18 Comparison Evaluation; Number of values to compute concerning the scheduling process, $(n > k)$=3

the evaluation method required to compute at least (105) during the optimisation process and (5) values during the assessment operation. Moreover, when the D term increased to $d = 3$, as shown in Figure5.18, and to $d = 4$, as shown in Figure 5.19, the difference in the computation complexity between the two operations increased noticeably, in comparison to the previous scenario where the D term is smaller d=2, as shown in Figure 5.17, and d=1, as shown in Figure 5.18. The D term had a greater influence on the computational complexity during the optimisation stage in comparison to the assessments stage. Furthermore, regardless of changes in the D term during the assessment stage, the computational complexity did not change over four separate scenarios. On the other hand, the I term influenced the computation process during the assessment stage. However, it is still relatively small in comparison to the optimisation.

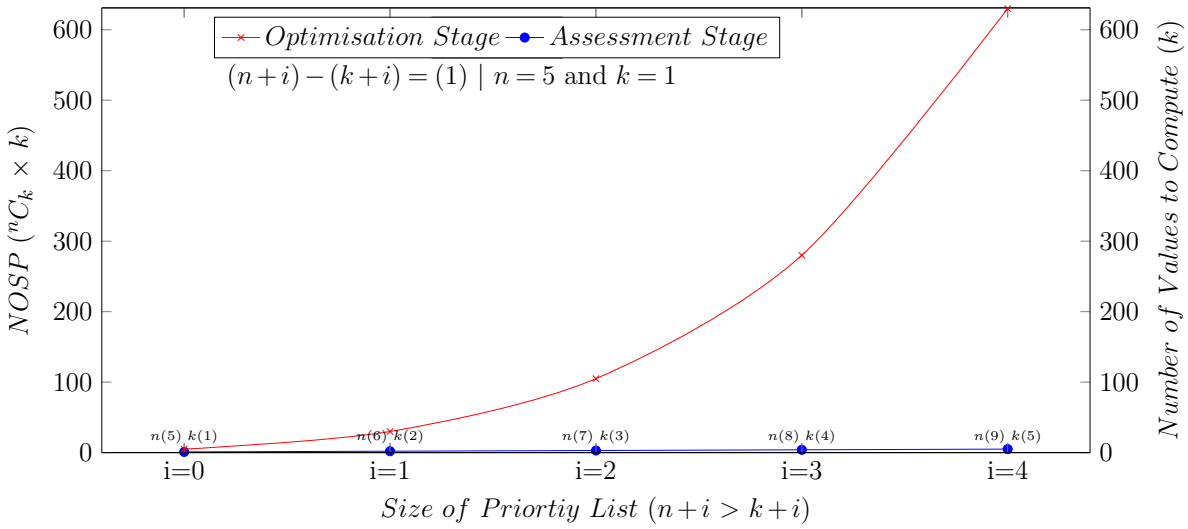


Fig. 5.19 Comparison Evaluation; Number of values to compute concerning the scheduling process,$(n > k)$=4

The second phase observes the computational complexity in response to the number of ranking instances. More directly, we intended to analyse the computational complexity in response to the increased number of k term, over four separate scenarios: first scenario k=1, second k=2,third k=3 and fourth k=4, where k is representing the number of desired events. Simultaneously, the real-time system is reporting an i^{th} number of undesired events for each scenario. Consequently, the size of the prioritised number of events n becomes $n \geqslant k$. In such cases, the number of ranking instances will increase systematically in response to the increased number of K and I inputs. Hence, the evaluation process intended to observe the computational complexity during the assessment stage and optimisation process, where the size of the priority list is $n = k+i$ for the underlying scenarios, as shown in Figure 5.20.

The computational complexity of the optimisation process has responded to the input of both terms i and k, as shown in Figure 5.21 part (a). However, during the assessment stage, the response was only to the k term as shown in Figure 5.21 part (b). Likewise, the computational complexity during the assessment stage is relatively very small in comparison to the optimisation process.

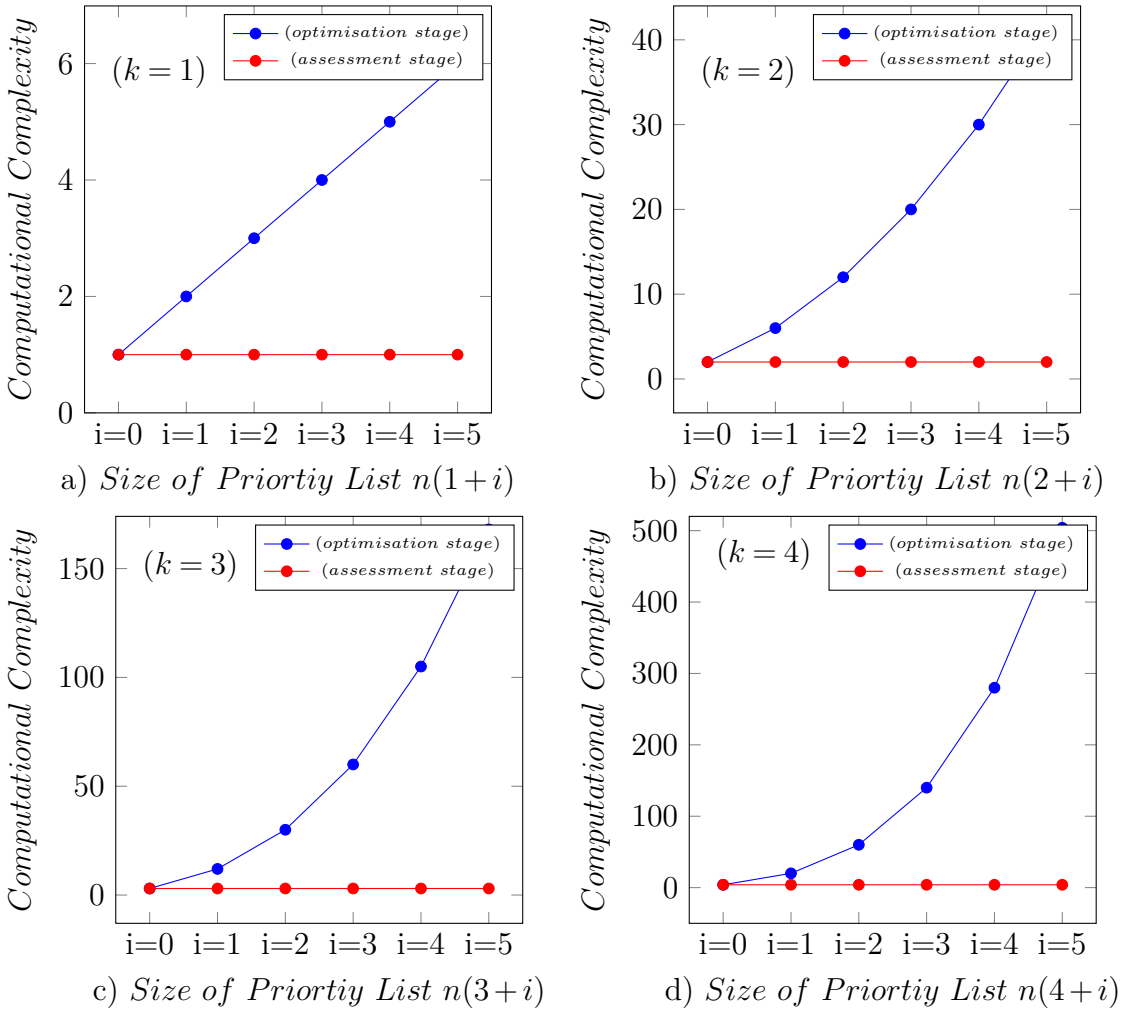


Fig. 5.20 The number of values required to be computed for various sizes of priority lists, concerning the scheduling process

We can conclude during the assessment stage, the employed method of evaluation is required to compute a lesser number of values in comparison to the optimisation process. Likewise, in the assessment process, the performance metric is required to compute only the necessary values for assessing the scheduling process of the real-time system, while the optimised operation is required to compute the values for all ranking instances obtained by the combination (${}^{n}C_{k}$) operation.

This section has explained the computational complexity concerning the scheduling process of the real-time system; the next section discusses our conclusion.

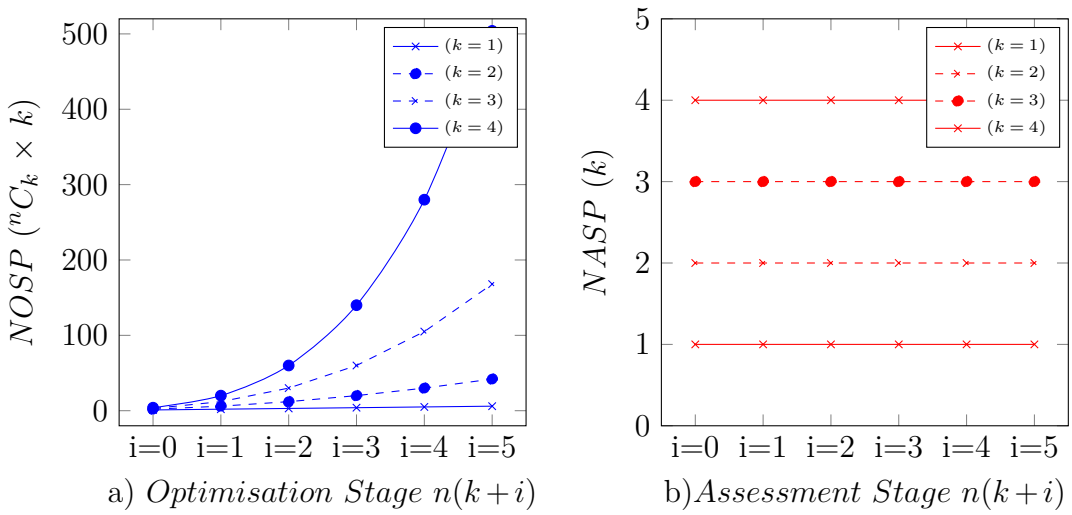


Fig. 5.21 The number of values it required to be computed for various sizes of priority lists, concerning the scheduling process

5.8 Conclusion

This chapter has introduced the Scheduling Capability Score for evaluating the scheduling process of the real time system.

The proposed performance metric has been validated with two levels of assessments. The first level was a case study based evaluation; this was to guide researchers from various disciplines on how to adopt the proposed performance metrics into their domain specific configuration. The second validation stage involved conducting a quality based evaluation; this was to examine the proposed performance against its intended purpose. More directly, the evaluation process examined the proposed performance metric and existing scoring scheme over three separate scenarios; the results obtained showed that the *Activity of Interest Score* did not provide a unique scoring scheme for all the ranking instances concerning the scheduling process of the real time system; the proposed performance metric accomplished this from the good to the worst state respectively and over three separate scenarios. Indeed, the proposed scoring scheme is

capable of notifying the analyst team about any ranking capability issues which may occur during a real time operation.

Finally, we have developed a method to analyse the computational complexity of the prioritisation and scheduling processes respectively. More directly, the proposed method of evaluation has discussed the computational complexity concerning the number of values required for completion during two different operations. The first operation was the assessment stage where the evaluation method was required to compute only the necessary values for assessing the ranking capability of the real time system. However, the second operation, during the optimisation process, required the evaluation method to compute more values in order to assess all the ranking instances for any given scenarios.

Chapter 6

Conclusion

Sea, land, space, air and cyberspaces are the five classifications of domains. It is essential that the analyst's team keep up to date with these in a dynamically monitored environment. To meet the needs of these domains, the data fusion community has introduced the information fusion reference model for multi-disciplinary areas; this describes the theoretical concept of a real time system with multiple levels of situational assessment. Each level performs a contextual task during a real time operation. In return, the proposed outputs of these simultaneous processes are either prioritised in a list of events (namely the tracking activity) or represented by visual means, supporting timely responses during a real time operation.

Usually, the capability of a SWA system, in terms of ranking the identified tracking activity, is hindered by the knowledge limitation problem, specifically when the underlying system is processing multiple sensor information during the real time operation. Consequently, the system may not rank the identified list of tracking activities as perfect, as desired by the decision-making resources. With this in mind, the researchers defined two further levels for assessing the real time system performance. The first level is the process refinement (level 4) for evaluating the information perception,

comprehension and projection process. The second level is the user refinement (level 5) for addressing the knowledge representation issues concerning the user-system relation.

The process refinement(Level 4) of the Joint Directors of Laboratories (JDL) is a meta-process used to assess and improve the data fusion task during a real time operation. During the assessment stage, the underlying process is expected to verify the SWA domains(perception, comprehension and projection) and it can take two forms of evaluation; qualitative and quantitative. Furthermore, during the qualitative stage, the evaluation process is required to have predefined knowledge in order to serve only a domain specific configuration. However, during the quantitative assessment, the evaluation method is capable of assessing different domains with minimum details or no predefined knowledge about domain specific information. Such methods can easily be applied to different domains and serve a wider number of systems in comparison to the qualitative method.

This thesis has developed advanced methods for evaluating the ranking capability of an SWA system, using an analytical approach for measuring the ranking capability in real time operations. The proposed scoring scheme can provide another dimensional support for the decision-making resources, specifically when the real time system is experiencing ranking capability issues.

Furthermore, the thesis has developed four different performance metrics using quantitative approaches for evaluating the ranking capability in a real time system. The *Ranking Capability Score RCS* has been designed to evaluate the prioritisation process for the SWA domains (perception, comprehension and projection). The *Scheduling Capability Score SCS* is used for evaluating the scheduling process for data fusion information again for the SWA domains. The *Ranking Capability Score* RCS' and the *Scheduling Capability Score'* SCS' were developed to address the knowledge representation problem concerning the user-system relation.

Additionally, this work has presented validating techniques to challenge the proposed scoring scheme with three levels of assessments. The first level is a case study based scenario. This is to guide researchers from different areas in adopting the proposed metrics to their domain specific needs and configuration.

The second level is a reliability-based assessment used to validate the proposed scoring scheme against the decision-making perception. The validation process attempted to verify the knowledge representation problem between the proposed scoring scheme and user perception in regards to the three following qualitative states. The best state is when the tracking activities are prioritised and scheduled entirely as perfect as the ground truth, the degraded state is when the proposed list of tracking activities are ranked as not perfect as the ground truth and, finally, the worst state is when the emerging events are ranked opposite to the ground truth.

The third level is a quality-based assessment for verifying the proposed performance metrics against their intended purpose. The underpinning evaluation encompassed three phases. The first phase used an analytical approach to compute the number of ranking instances for any given scenario concerning the prioritisation process or scheduling process of a real time system. The second phase used Matlab to simulate all ranking instances being computed during the analytical stage. Finally, the third phase examined the potential proposed scoring scheme in terms of providing a unique score for all possible ranking instances being proposed by the simulation phase.

The proposed performance metrics were designed and evaluated using an analytical approach. Such methods allowed the evaluation process to conduct a rigorous analysis of the prioritisation and scheduling processes, despite any constraints related to a domain-specific configuration.

Moreover, the work developed a method to analyse the computational complexity for two different operations involved in evaluating the prioritisation and scheduling

processes for a real time system. The first operation is during the assessment stage, where the underlying performance metric is required to compute only essential values for assessing the capability of a real time system. The second operation occurs when the potential performance metric is required to compute more values to assess the optimisation technique, concerning the ranking capability of the SWA system.

The proposed performance metrics have been designed for assessing the information fusion on perception, comprehension and projection, using quantitative assessments method to serve a wider number of domains. For future work, the proposed solution can be extended as follows:

The first direction; the proposed performance metric can be used for a domain specific scenario, not only for assessing the ranking capability but also to act as an automated optimisation process for sensing the three following operations: the prioritisation paradigms for any desired tracking activity, the scheduling paradigms for different classes of tracking activities and sensing the different scheduling and prioritisation paradigms for time sensitive operations.

The second direction; this project has developed two separate performance metrics for assessing the prioritisation and scheduling processes during real time operations. Future work can involve a combined approach of the two.

References

[1] (2016). Oxford learner's dictionaries.

[2] Alberto, Roman, L., Dominique, K., and Julio, C. (2013). *OSSIM Open Source Security Information Management*.

[3] AlientVault (2014). *Alient Vault*.

[4] Barford, P., Dacier, M., Dietterich, T., Fredrikson, M., Giffin, J., Jajodia, S., Jha, S., Li, J., Liu, P., Ning, P., et al. (2010). Cyber sa: Situational awareness for cyber defense. *Cyber Situational Awareness*, pages 3–13.

[5] Bhatt, P., Yano, E. T., Amorim, J., and Gustavsson, P. (2014). A cyber security situational awareness framework to track and project multistage cyber attacks. In *Proceedings of the 9th International Conference on Cyber Warfare & Security: ICCWS 2014*, page 356. Academic Conferences Limited.

[6] Bier, V. M., Haphuriwat, N., Menoyo, J., Zimmerman, R., and Culpen, A. M. (2008). Optimal resource allocation for defense of targets based on differing measures of attractiveness. *Risk Analysis*, 28(3):763–770.

[7] Blasch, E., Breton, R., Valin, P., and Bosse, E. (2011a). User information fusion decision making analysis with the c-ooda model. In *Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on*, pages 1–8.

[8] Blasch, E. and Plano, S. (2005a). Dfig level 5 (user refinement) issues supporting situational assessment reasoning. In *Information Fusion, 2005 8th International Conference on*, volume 1, pages xxxv–xliii. IEEE.

[9] Blasch, E. and Plano, S. (2005b). Proactive decision fusion for site security. In *Information Fusion, 2005 8th International Conference on*, volume 2, pages 8 pp.–.

[10] Blasch, E., Steinberg, A., Das, S., Llinas, J., Chong, C., Kessler, O., Waltz, E., and White, F. (2013). Revisiting the jdl model for information exploitation. In *Information Fusion (FUSION), 2013 16th International Conference on*, pages 129–136. IEEE.

[11] Blasch, E., Valin, P., Bosse, E., Nilsson, M., van Laere, J., and Shahbazian, E. (2009). Implication of culture: user roles in information fusion for enhanced situational understanding. In *Information Fusion, 2009. FUSION'09. 12th International Conference on*, pages 1272–1279. IEEE.

[12] Blasch, E. P. (2003). Situation, impact, and user refinement. In *AeroSense 2003*, pages 463–472. International Society for Optics and Photonics.

[13] Blasch, E. P., Breton, R., and Valin, P. (2011b). Using the c-ooda model for cimic analysis. In *Aerospace and Electronics Conference (NAECON), Proceedings of the 2011 IEEE National*, pages 130–138. IEEE.

[14] Blasch, E. P., Rogers, S. K., Holloway, H., Tierno, J., Jones, E. K., and Hammoud, R. I. (2014). Quest for information fusion in multimedia reports. *International Journal of Monitoring and Surveillance Technologies Research (IJMSTR)*, 2(3):1–30.

[15] Blasch, E. P., Salerno, J. J., and Tadda, G. P. (2011c). Measuring the worthiness of situation assessment. In *Aerospace and Electronics Conference (NAECON), Proceedings of the 2011 IEEE National*, pages 87–94. IEEE.

[16] Bonelli, L., Antonio, L., Luisa, G., and Andrea, M. (2013). *logging Managment for Storage Cloud Compliance*.

[17] Boyd, J. R. (1987). *Destruction and creation*. US Army Comand and General Staff College.

[18] Boyd, J. R. (1995). The essence of winning and losing, 28 june 1995.

[19] Breton, R., Bosse, E., Rousseau, R., and Tremblay, S. (2012). Framework for the analysis of information relevance (fair). In *Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), 2012 IEEE International Multi-Disciplinary Conference on*, pages 210–213.

[20] Breton, R. and Rousseau, R. (2005). The c-ooda: A cognitive version of the ooda loop to represent c2 activities. In *Proceedings of the 10th International Command and Control Research Technology Symposium*.

[21] Byers, S. R. and Yang, S. J. (2008). Real-time fusion and projection of network intrusion activity. In *Information Fusion, 2008 11th International Conference on*, pages 1–8. IEEE.

[22] Chien, S. and Ho, C. (2012). A novel threat prediction framework for network security. *Advances in Information Technology and Industry Applications*, pages 1–9.

[23] CSRS-Corop (2014). *Cyber Security Research and Solutions Corporation*.

[24] Dahlbom, A. (2011). *Petri nets for situation recognition*. PhD thesis, Ã–rebro University, School of Science and Technology. Anders Dahlbom is also affiliated to SkÃ¶vde Artificial Intelligence Lab (SAIL), Information Fusion Research Program, HÃ¶gskolan i SkÃ¶vde .

[25] Du, H., Liu, D. F., Holsopple, J., and Yang, S. J. (2010). Toward ensemble characterization and projection of multistage cyber attacks. In *Computer Communications and Networks (ICCCN), 2010 Proceedings of 19th International Conference on*, pages 1–8. IEEE.

[26] Endsley, M. (2000). Theoretical underpinnings of situation awareness: A critical review. *Situation awareness analysis and measurement*, pages 3–32.

[27] Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1):32–64.

[28] Epiphaniou, G., French, T., and Maple, C. (2014). The dark web: Cyber-security intelligence gathering opportunities, risks and rewards. *CIT. Journal of Computing and Information Technology*, 22(LISS 2013):21–30.

[29] Farhadi, H., AmirHaeri, M., and Khansari, M. (2011). Alert correlation and prediction using data mining and hmm. *The ISC Journal of Information Security*, 3:77 – 101.

[30] Fava, D., Holsopple, J., JayYang, S., and Argauer, B. (2007). Terrain and behavior modeling for projecting multistage cyber attacks. In *Information Fusion, 2007 10th International Conference on*, pages 1–7. IEEE.

[31] Fava, D. S., Byers, S. R., and Yang, S. J. (2008). Projecting cyberattacks through variable-length markov models. *Information Forensics and Security, IEEE Transactions on*, 3(3):359–369.

[32] Fischer, F. and Keim, D. A. (2014). Nstreamaware: real-time visual analytics for data streams to enhance situational awareness. In *Proceedings of the Eleventh Workshop on Visualization for Cyber Security*, pages 65–72. ACM.

[33] Foo, P. H. and Ng, G. W. (2013). High-level information fusion: An overview. *J. Adv. Inf. Fusion*, 8(1):33–72.

[34] Giacobe, N. A. (2010). Application of the jdl data fusion process model for cyber security. In *SPIE Defense, Security, and Sensing*, pages 77100R–77100R. International Society for Optics and Photonics.

[35] Giacobe, N. A. (2013). *Measuring the effectiveness of visual analytics and data fusion techniques on situation awareness in cyber-security*. PhD thesis, The Pennsylvania State University.

[36] Grant, T., Venter, H., and Eloff, J. (2007). Simulating adversarial interactions between intruders and system administrators using ooda-rr. In *Proceedings of the 2007 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries*, pages 46–55. ACM.

[37] Gross, G. A., Schlegel, D. R., Corso, J. J., Llinas, J., Nagi, R., Shapiro, S. C., et al. (2014). Systemic test and evaluation of a hard+ soft information fusion framework: Challenges and current approaches. In *Information Fusion (FUSION), 2014 17th International Conference on*, pages 1–8. IEEE.

[38] Hausken, K. (2017). Defense and attack for interdependent systems. *European Journal of Operational Research*, 256(2):582–591.

[39] Hausken, K. and He, F. (2016). On the effectiveness of security countermeasures for critical infrastructures. *Risk Analysis*, 36(4):711–726.

[40] Holsopple, J., Sudit, M., Nusinov, M., Liu, D., Du, H., and Yang, S. (2010). Enhancing situation awareness via automated situation assessment. *Communications Magazine, IEEE*, 48(3):146–152.

[41] Holsopple, J., Yang, S., and Sudit, M. (2006). Tandi: threat assessment of network data and information. SPIE.

[42] Holsopple, J. and Yang, S. J. (2008). Fusia: Future situation and impact awareness. In *Information Fusion, 2008 11th International Conference on*, pages 1–8. IEEE.

[43] Howard, N. and Kanareykin, S. (2012a). The case for intention awareness in security systems. *international Conference on intelligent Computational Systems*, pages 91 – 95.

[44] Howard, N. and Kanareykin, S. (2012b). Intention awareness in cyber security. In *Cyber Security, Cyber Warfare and Digital Forensic (CyberSec), 2012 International Conference on*, pages 6–11. IEEE.

[45] Hughes, J. (2013). Top 5 open source event correlation tools.

[46] Hutchins, E. M., Cloppert, M. J., and Amin, R. M. (2011). Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. *Leading Issues in Information Warfare & Security Research*, 1:80.

[47] Ingols, K., Chu, M., Lippmann, R., Webster, S., and Boyer, S. (2009). Modeling modern network attacks and countermeasures using attack graphs. In *Computer Security Applications Conference, 2009. ACSAC'09. Annual*, pages 117–126. IEEE.

[48] Jajodia, S. and Noel, S. (2010). Advanced cyber attack modeling analysis and visualization. Technical report, DTIC Document.

[49] Kim, H., Kim, S., and Kim, S. (2012). Decision support system for zero-day attack response. *Appl. Math*, 6(1S):221S–241S.

[50] Klein, G. A. (1993). A recognition-primed decision (rpd) model of rapid decision making. *Decision making in action: Models and methods*, 5(4):138–147.

[51] Lambert, D. A. (2007). STDF model based maritime situation assessments. In *International Conference on Information Fusion*.

[52] Lathrop, S., Hill, J., and Surdu, J. (2003). Modeling network attacks. In *Proceedings of the 12th Conference on Behavior Representation in Modeling and Simulation*, pages 401–407.

[53] Llinas, J., Bowman, C., Rogova, G., Steinberg, A., Waltz, E., and White, F. (2004). Revisiting the jdl data fusion model ii. Technical report, DTIC Document.

[54] Llinas, J., Liggins, M. E., and Hall, D. L. (2008). *Handbook of Multisensor Data Fusion: Theory and Practice*. CRC Press.

[55] Llinas, J., Snidaro, L., García, J., and Blasch, E. (2016). Context and fusion: Definitions, terminology. In *Context-Enhanced Information Fusion*, pages 3–23. Springer.

[56] Mahboubian, M., Udzir, N., Subramaniam, S., and Hamid, N. (2012). An alert fusion model inspired by artificial immune system. In *Cyber Security, Cyber Warfare and Digital Forensic (CyberSec), 2012 International Conference on*, pages 317–322. IEEE.

[57] Mathew, S., Britt, D., Giomundo, R., Upadhyaya, S., Sudit, M., and Stotz, A. (Oct.). Real-time multistage attack awareness through enhanced intrusion alert clustering. In *Military Communications Conference, 2005. MILCOM 2005. IEEE*, pages 1801–1806 Vol. 3.

[58] Mathew, S., Shah, C., and Upadhyaya, S. (2005). An alert fusion framework for situation awareness of coordinated multistage attacks. In *Information Assurance, 2005. Proceedings. Third IEEE International Workshop on*, pages 95 – 104.

[59] Maughan, D. (2009). National cyber security research assessment and roadmap. In *Proceedings of the 5th Annual Workshop on Cyber Security and Information Intelligence Research: Cyber Security and Information Intelligence Challenges and Strategies*, CSIIRW '09, pages 6:1–6:1, New York, NY, USA. ACM.

[60] McAfeeLab (2012). 2012 threat predictions. Report, McAfee An Intel Company.

[61] McNeese, M., Cooke, N., Dâ€™Amico, A., Endsley, M., Gonzalez, C., Roth, E., and Salas, E. (2012). Perspectives on the role of cognition in cyber security. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 56, pages 268–271. SAGE Publications.

[62] Morris, T., Mayron, L., Smith, W., Knepper, M., Ita, R., and Fox, K. (2011). A perceptually-relevant model-based cyber threat prediction method for enterprise mission assurance. In *Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), 2011 IEEE First International Multi-Disciplinary Conference on*, pages 60–65. IEEE.

[63] Nicol, D., Sanders, W., and Trivedi, K. (2004). Model-based evaluation: From dependability to security. *Dependable and Secure Computing, IEEE Transactions on*, 1(1):48–65.

[64] Noel, S. and Jajodia, S. (2008). Optimal ids sensor placement and alert prioritization using attack graphs. *Journal of Network and Systems Management*, 16(3):259–275.

[65] Patterson, S. A. and Apostolakis, G. E. (2007). Identification of critical locations across multiple infrastructures for terrorist actions. *Reliability Engineering & System Safety*, 92(9):1183–1203.

[66] Raulerson, E. L. (2013). Modeling cyber situational awareness through data fusion. Technical report, DTIC Document.

[67] Raulerson, E. L., Hopkinson, K. M., and Laviers, K. R. (2014). A framework to facilitate cyber defense situational awareness modeled in an emulated virtual machine testbed. *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, page 1548512914552530.

[68] Rebovich Jr, G. (2005). Enterprise systems engineering theory and practice, vol. 2: Systems thinking for the enterprise: New and emerging perspectives. *The MITRE Corporation, MP05B0000043*.

[69] Rouillard, J. P. (2004). Real-time log file analysis using the simple event correlator (sec). In *LISA*, pages 133–150.

[70] Roy, A., Kim, D., and Trivedi, K. (2011). Attack countermeasure trees (act): towards unifying the constructs of attack and defense trees. *Security and Communication Networks*.

[71] Roy, A., Kim, D., and Trivedi, K. (2012). Scalable optimal countermeasure selection using implicit enumeration on attack countermeasure trees. In *Dependable Systems and Networks (DSN), 2012 42nd Annual IEEE/IFIP International Conference on*, pages 1–12. IEEE.

[72] Roy, J. (2001). From data fusion to situation analysis. In *Proceedings of the Fourth International Conference on Information Fusion (FUSION 2001), Montreal, Canada*.

[73] Salerno, J. (2007). Where's level 2/3 fusion-a look back over the past 10 years. In *Information Fusion, 2007 10th International Conference on*, pages 1–4. IEEE.

[74] Salerno, J. (2008). Measuring situation assessment performance through the activities of interest score. In *Information Fusion, 2008 11th International Conference on*, pages 1 –8.

[75] Salerno, J., Sudit, M., Yang, S., Tadda, G., Kadar, I., and Holsopple, J. (2010). Issues and challenges in higher level fusion: Threat/impact assessment and intent modeling (a panel summary). In *Information Fusion (FUSION), 2010 13th Conference on*, pages 1 –17.

[76] Salerno, J. J., Blasch, E. P., Hinman, M., and Boulware, D. M. (2005). Evaluating algorithmic techniques in supporting situation awareness. In *Defense and Security*, pages 96–104. International Society for Optics and Photonics.

[77] Salerno, J. J. and Tadda, G. P. (2009). Ranking activities based on their impact and threat. Technical report, DTIC Document.

[78] Shurrab, O. (2016a). Responding to emerging situation by developing the ranking capability score (rcs). In *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2012 IEEE Conference on*, pages 582–587. IEEE.

[79] Shurrab, O. and Awan, I. (2015a). Measuring the ranking capability of swa system. In *Future Internet of Things and Cloud (FiCloud), 2015 3rd International Conference on*, pages 165–172.

[80] Shurrab, O. and Awan, I. (2015b). Performance evaluation for process refinement stage of swa system. In *Future Internet of Things and Cloud (FiCloud), 2015 3rd International Conference on*, pages 240–247. IEEE.

[81] Shurrab, O. M. (2016b). Toward an optimisation technique for dynamically monitored environment. In *SPIE Remote Sensing*, pages 100070G–100070G. International Society for Optics and Photonics.

[82] Steinberg, A. and Bowman, C. (2004). Rethinking the jdl data fusion levels. *NSSDF JHAPL*.

[83] Steinberg, A., Bowman, C., and White, F. (1998). *Revisions to the JDL data fusion model*. American inst of aeronautics and astronautics new york.

[84] Steinberg, A. N. and Bowman (1999). Revisions to the jdl data fusion model. *International Society for Optics and Photonics*, pages 430–441.

[85] Stotz, A. and Sudit, M. (2007). Information fusion engine for real-time decision-making (inferd): A perceptual system for cyber attack tracking. In *Information Fusion, 2007 10th International Conference on*, pages 1 –8.

[86] Sudit, M., Holender, M., Stotz, A., Rickard, T., and Yager, R. (2007). Inferd and entropy for situational awareness. *Journal Adv. Info. Fusion*, 2(1).

[87] Sudit, M., Stotz, A., and Holender, M. (2005a). Situational awareness of a coordinated cyber attack. In *Proc. SPIE*, volume 5812, pages 114–129.

[88] Sudit, M., Yang, S., Kuhl, M., Stotz, A., Holender, M., Holsopple, J., Bohannon, E., and Kistner, J. (2005b). High level fusion in the cyber domain. Technical report, DTIC Document.

[89] Tadda, G. (2008). Measuring performance of cyber situation awareness systems. In *Information Fusion, 2008 11th International Conference on*, pages 1 –8.

[90] Tadda, G. and Salerno, J. (2010). Overview of cyber situation awareness. *Cyber Situational Awareness*, pages 15–35.

[91] Vaarandi, R. (2002). Sec-a lightweight event correlation tool. In *IP Operations and Management, 2002 IEEE Workshop on*, pages 111–115. IEEE.

[92] Vaarandi, R. and Grimaila, M. R. (2012). Security event processing with simple event correlator. *ISSA Jounal*, pages 30–37.

[93] VALIN, P. (2013). Potential information fusion technologies applicable to maritime piracy awareness. *Prediction and Recognition of Piracy Efforts Using Collaborative Human-Centric Information Systems*, 109:117.

[94] WebsenseLabs (2012). Security predictions for 2012. Report, Websense security Lab.

[95] Wohl, J. G. (1981). Force management decision requirements for air force tactical command and control. *Systems, Man and Cybernetics, IEEE Transactions on*, 11(9):618–639.

[96] Wu, Q., Zheng, R., Li, G., and Zhang, J. (2011). Intrusion intention identification methods based on dynamic bayesian networks. *Procedia Engineering*, 15:3433–3438.

[97] Yang, S. J., Holsopple, J., and Liu, D. (2009a). Elements of impact assessment: a case study with cyber attacks. In *SPIE Defense, Security, and Sensing*, pages 73520D–73520D. International Society for Optics and Photonics.

[98] Yang, S. J., Stotz, A., Holsopple, J., Sudit, M., and Kuhl, M. (2009b). High level information fusion for tracking and projection of multistage cyber attacks. *Information Fusion*, 10(1):107–121.

[99] Zhang, H., Shi, J., and Chen, X. (2013). A multi-level analysis framework in network security situation awareness. *Procedia Computer Science*, 17:530–536.